

#2

日 本 国 特 許 庁
JAPAN PATENT OFFICE

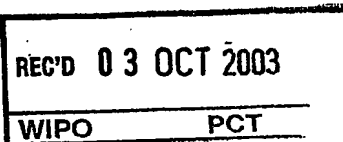
PCT/JP 03/10309

13.08.03

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2002年 8月14日



出 願 番 号
Application Number: 特願2002-236292
[ST. 10/C]: [JP 2002-236292]

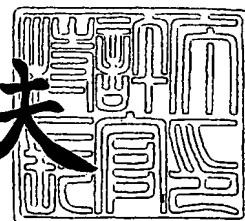
出 願 人
Applicant(s): 独立行政法人産業技術総合研究所
アマシャム バイオサイエンス株式会社

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2003年 9月19日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



BEST AVAILABLE COPY

【書類名】 特許願
【整理番号】 021634
【提出日】 平成14年 8月14日
【あて先】 特許庁長官殿
【国際特許分類】 C12M
C07H

【発明者】

【住所又は居所】 茨城県つくば市梅園 1 - 1 - 1 独立行政法人産業技術
総合研究所内

【氏名】 成松 久

【発明者】

【住所又は居所】 茨城県つくば市梅園 1 - 1 - 1 独立行政法人産業技術
総合研究所内

【氏名】 後藤 雅式

【発明者】

【住所又は居所】 茨城県つくば市梅園 1 - 1 - 1 独立行政法人産業技術
総合研究所内

【氏名】 佐藤 隆

【特許出願人】

【識別番号】 301021533

【氏名又は名称】 独立行政法人産業技術総合研究所

【代表者】 理事長 吉川 弘之

【特許出願人】

【識別番号】 597145779

【氏名又は名称】 アマシャム バイオサイエンス株式会社

【代理人】

【識別番号】 100089705

【住所又は居所】 東京都千代田区大手町二丁目 2 番 1 号 新大手町ビル 2
0 6 区 ユアサハラ法律特許事務所

【弁理士】

【氏名又は名称】 社本 一夫

【電話番号】 03-3270-6641

【選任した代理人】

【識別番号】 100076691

【弁理士】

【氏名又は名称】 増井 忠次

【選任した代理人】

【識別番号】 100075270

【弁理士】

【氏名又は名称】 小林 泰

【選任した代理人】

【識別番号】 100080137

【弁理士】

【氏名又は名称】 千葉 昭男

【選任した代理人】

【識別番号】 100096013

【弁理士】

【氏名又は名称】 富田 博行

【選任した代理人】

【識別番号】 100107386

【弁理士】

【氏名又は名称】 泉谷 玲子

【手数料の表示】

【予納台帳番号】 051806

【納付金額】 10,500円

【その他】 国等以外のすべての者の持分の割合 50/100。国等の委託研究の成果に係わる特許出願（平成13年度新エネルギー・産業技術総合開発機構 糖鎖合成関連遺伝子ライブラリーの構築委託研究，産業活力再生特別措置法第30条の適用を受けるもの）。

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 新規N-アセチルガラクトサミン転移酵素およびそれをコードする核酸

【特許請求の範囲】

【請求項1】 配列番号1若しくは3に示されるアミノ酸配列、あるいは該アミノ酸配列において1若しくは複数個のアミノ酸が置換し若しくは欠失し、または該アミノ酸配列に1若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有し、N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する活性を有するタンパク質。

【請求項2】 前記タンパク質が、配列番号1に示されるアミノ酸配列と50%以上の同一性を有する、請求項1記載のタンパク質。

【請求項3】 前記タンパク質が、配列番号1に示されるアミノ酸配列と60%以上の同一性を有する、請求項1記載のタンパク質。

【請求項4】 請求項1ないし3のいずれか1項に記載のタンパク質をコードする核酸。

【請求項5】 配列番号2または4に示される塩基配列を有する核酸とストリンジントな条件下でハイブリダイズする、請求項1記載のタンパク質をコードする核酸。

【請求項6】 配列番号2に示される塩基配列のヌクレオチド1-3120または配列番号4に示される塩基配列のヌクレオチド1-2997を有する、請求項6記載の核酸。

【請求項7】 請求項4ないし6のいずれか1項に記載の核酸を含み、宿主細胞中で該核酸を発現することができる組換えベクター。

【請求項8】 請求項7の組換えベクターで形質転換されている宿主細胞。

【請求項9】 請求項4記載の核酸とストリンジントな条件下でハイブリダイズする測定用核酸。

【請求項10】 前記測定用核酸がプライマーとして使用される、配列番号20、21、23および24からなる群から選択される、請求項9記載の測定用核酸。

【請求項 11】 前記測定用核酸がプローブとして使用される、配列番号 2 または 25 である、請求項 9 記載の測定用核酸。

【請求項 12】 前記測定用核酸が癌マーカーとして使用される、請求項 9 記載の測定用核酸。

【請求項 13】 請求項 10 ないし 12 のいずれか 1 項に記載の測定用核酸、および測定説明書を含む、測定用キット。

【請求項 14】 請求項 1 ないし 5 のいずれか 1 項に記載のタンパク質に結合する抗体。

【請求項 15】 モノクローナル抗体である、請求項 14 記載の抗体。

【請求項 16】 生物試料の癌化を検定する方法であって、

(a) 生物試料中の請求項 1 ないし 5 のいずれか 1 項に記載のタンパク質を定量し；そして

(b) 生物試料中の前記タンパク質の定量値が、対照の生物試料中の前記タンパク質の定量値よりも大きい場合には癌化していると判断する工程を含む、前記方法。

【請求項 17】 請求項 15 または 16 に記載の抗体を用いてタンパク質を定量する、請求項 16 記載の検定方法。

【請求項 18】 生物試料の癌化を検定する方法であって、

(a) 生物試料中の請求項 4 記載の核酸を定量し；そして

(b) 生物試料中の請求項 4 記載の核酸の定量値が、対照の生物試料中の前記核酸の定量値の 1.5 倍以上である場合には癌化していると判断する工程を含む、前記方法。

【請求項 19】 (a) 生物試料中の請求項 4 記載の核酸と少なくとも 1 つの請求項 9 記載の測定用核酸とをハイブリダイズさせ；

(b) 請求項 4 記載の核酸を増幅させ；

(c) 増幅産物に請求項 9 記載の測定用核酸をハイブリダイズさせ；

(d) 前記増幅産物とハイブリダイズした前記測定用核酸から生じるシグナルを定量し；そして

(e) 前記シグナルの定量値が、対照の生物試料中の対応するシグナルの定量

値の 1.5 倍以上である場合には癌化していると判断する工程を含む、請求項 18 記載の方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する活性を有する新規な酵素およびそれをコードする核酸、並びに該核酸を測定するための核酸に関する。

【0002】

【従来の技術】

様々な生物において、N-アセチルガラクトサミン-N-アセチルグルコサミンの二糖のリンケージを有する構造が、糖タンパク質や糖脂質の糖鎖中に見出されている[参考文献1、2を参照]。ヒトでは、この二糖構造は β 1-4リンケージ(GalNAc β 1-4GlcNAc)として知られており、N-グリカン中にのみ見出されている[参考文献3を参照]。本構造を含むヒト型糖鎖を入手するには、煩雑な化学合成を行うか、あるいは天然のタンパク質から分取する方法しかない。さらに、生体内では、上記二糖構造において、N-アセチルガラクトサミンの代わりにガラクトースであるものが含まれるため、標的とする二糖構造を有する糖鎖を分取する場合に非常に労力を要する。

【0003】

本出願前に、本発明者らは、グルクロン酸やポリペプチドにN-アセチルガラクトサミンを転移する活性を有する酵素として、ppGalNAc-T10、-T11、-T12、-T13、-T14、-T15、-T16、-T17、CSGalNAc-T1、および-T2を同定し、さらにこれらの遺伝子構造についても明らかにしている。これまで、上記N-アセチルガラクトサミンを転移する活性を有するN-アセチルガラクトサミン転移酵素は、少なくとも22種類知られており(表1)、各々、受容体基質特異性が異なる。

【0004】

【表 1】

表 1 N-アセチルガラクトサミン転移酵素とその基質特異性

正式名	略称	由来	基質特異性	文献
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] I	ppGalNAc-T1	ヒト	Ser/Thr	White, T. 5 (1995)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] II	ppGalNAc-T2	ヒト	Ser/Thr	White, T. 5 (1995)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] III	ppGalNAc-T3	ヒト	Ser/Thr	Bennet, E. P. 5 (1996)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] IV	ppGalNAc-T4	ヒト	Ser/Thr	Bennet, E. P. 5 (1998)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] V	ppGalNAc-T6	ヒト	Ser/Thr	Bennet, E. P. 5 (1999) (1)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] VI	ppGalNAc-T7	ヒト	Ser/Thr	Bennet, E. P. 5 (1999) (2)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] VII	ppGalNAc-T8	ヒト	Ser/Thr	White, K. E. 5 (2000)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] VIII	ppGalNAc-T9	ヒト	Ser/Thr	Toba, S. 5 (2000)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] IX	ppGalNAc-T10	ヒト	Ser/Thr	特願 2001-401455 (未公表)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] X	ppGalNAc-T11	ヒト	Ser/Thr	特願 2001-401507 (未公表)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] XI	ppGalNAc-T12	ヒト	Ser/Thr	特願 2001-401507 (未公表)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] XII	ppGalNAc-T13	ヒト	Ser/Thr	特願 2001-401507 (未公表)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] XIII	ppGalNAc-T14	ヒト	Ser/Thr	Guo, J. M. 5 (2002)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] XIV	ppGalNAc-T15	ヒト	Ser/Thr	特願 2001-401507 (未公表)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] XV	ppGalNAc-T16	ヒト	Ser/Thr	特願 2001-401507 (未公表)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] XVI	ppGalNAc-T17	ヒト	Ser/Thr	特願 2001-401507 (未公表)
UDP-GalNAc:ド・リヘ・アチト・N-アセチル・ラクチシニトランスフェラーゼ [*] XVII	β 4GalNAcT	ヒト	GM3, GD3, LacCer	Nagata, Y. 5 (1992)
β 1, 4-N-アセチル・ラクチシニトランスフェラーゼ [*]				
UDP-GalNAc:H-α 1, 3-N-アセチル・ラクチシニトランスフェラーゼ [*]	ヒト血 A 群トランスフェラーゼ [*]	ヒト	Fuc α 1, 2Gal β 1-R	Yamamoto, F. 5 (1990)
UDP-GalNAc:ド・リヘ・アチト・α 1, 3-N-アセチル・ラクチシニトランスフェラーゼ [*] I	ホスホ糖脂質シクター [*]	ヒト	GalNAc β 1-3Gal α 1	Xu, H. 5 (1999)
UDP-GalNAc:ド・リヘ・アチト・α 1, 3-N-アセチル・ラクチシニトランスフェラーゼ [*] I			-4Gal β 1-3Glc-Car	
UDP-GalNAc:ド・リヘ・アチト・α 1, 3-N-アセチル・ラクチシニトランスフェラーゼ [*] I			GlcA	特願 2002-129156 (未公表)
UDP-GalNAc:ド・リヘ・アチト・α 1, 3-N-アセチル・ラクチシニトランスフェラーゼ [*] I			GlcA	特願 2002-24202 (未公表)

【0005】

【発明が解決しようとする課題】

N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する活性を有する酵素を単離し、その遺伝子の構造を明らかにすることにより、該酵素の遺伝

子工学的な生産や、該遺伝子に基づく疾患の診断が可能になる。しかしながら、該酵素は、未だ精製分離もされておらず、該酵素の単離および遺伝子の同定についての手がかりはない。そのために、該酵素に対する抗体も作製されていない。

【0006】

従って、本発明の目的は、N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する活性を有するタンパク質およびそれをコードする核酸を提供することである。また、本発明の目的は、該核酸を宿主細胞内で発現する組換えベクターおよび該核酸が導入され、前記核酸および該タンパク質を発現する細胞を提供することである。また、発現した該タンパク質は抗体作製用に利用することができる。従って、本発明の目的はまた、該タンパク質の生産方法を提供することでもある。さらに、発現したタンパク質および該タンパク質に対する抗体は、免疫組織染色、RIAおよびEIA等の免疫測定法に応用できる。さらにまた、本発明の目的は、上記本発明の核酸を測定するための測定用核酸を提供することでもある。

【0007】

【課題を解決するための手段】

上記の通り、目的とする酵素は、未だ単離されていないので、その部分アミノ酸配列を知ることもできない。一般に、細胞に微量しか含まれていないタンパク質を単離精製することは容易ではない。したがって、現在に至るまで単離されていない酵素を細胞から単離することは容易でないことが予想される。そこで、本発明者らは、目的とする酵素と比較的類似した作用を有する種々の酵素遺伝子の塩基配列間に、同一性が高いと考えられる領域を標的として、その相同配列を有すると思われる目的とする酵素を単離および同定することを試みた。具体的には、まず、公知の β 1, 4-ガラクトース転移酵素の塩基配列を検索し、相同領域を見出した。次に、この相同領域をもとにしてプライマーを設計し、cDNAライブラリーから5'RACE (rapid amplification of cDNA ends) 法で全長のオープンリーディングフレームを同定した。さらに、PCRで該酵素の遺伝子のクローニングに成功し、その塩基配列および推定アミノ酸配列を決定することによって、本発明を完成するに至った。

【0008】

本発明は、N-アセチルガラクトサミンを転移する活性を有するタンパク質およびこれらをコードする核酸を提供することにより、当該技術分野におけるこれらの多様な必要性を満たすのを援助する。

【0009】

即ち、本発明によれば、N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する活性を有するタンパク質が提供される。

本発明のタンパク質は、典型的には、配列番号2または4に記載の塩基配列から推定される、配列番号1または3に記載のアミノ酸配列を有する。

【0010】

本発明のタンパク質には、配列番号1または3に記載のアミノ酸配列のみでなく、当該配列と50%以上の同一性を有するものも含まれ、当該配列において1若しくは複数個のアミノ酸が置換し若しくは欠失し、または当該配列において1若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有するタンパク質が含まれる。

【0011】

本発明のタンパク質は、好ましくは、配列番号1または3に記載のアミノ酸配列と60%以上、より好ましくは70%以上、さらに好ましくは80%以上、さらになお好ましくは90%以上、最も好ましくは95%以上の同一性を有するアミノ酸配列を有する。

【0012】

本発明によれば、本発明のタンパク質をコードする核酸が提供される。

本発明の核酸は、典型的には、配列番号2または4に記載の塩基配列、前記塩基配列中に1から複数個の塩基の置換、欠失、挿入および／または付加を有する塩基配列、または前記塩基配列とストリンジェントな条件下でハイブリダイズする塩基配列を有し、および上記配列に相補的な核酸も含む。限定されるわけではないが、配列番号2に記載の塩基配列におけるヌクレオチド1-3120、および配列番号4に記載の塩基配列におけるヌクレオチド1-2997を含む核酸もまた本発明の一態様である。

【0013】

本発明によれば、本発明の核酸を含む組換えベクターが提供される。本発明の組換えベクターは、好ましくは発現ベクターである。

本発明によれば、本発明の組換えベクターを宿主細胞に導入して得られる形質転換体が提供される。

【0014】

本発明によれば、本発明のタンパク質をコードする核酸とストリンジェントな条件下でハイブリダイズする測定用核酸が提供される。本発明の測定用核酸をプライマーとして核酸増幅用反応に使用する場合には、配列番号20、21、23および24のいずれかに記載の配列を有することが好ましい。本発明の測定用核酸をプローブとして前記タンパク質をコードする核酸の測定に使用する場合には、配列番号22または24に記載の配列を有することが好ましい。さらに、本発明の測定用核酸は、癌マーカーとして使用することができる。

【0015】

本発明によれば、本発明の核酸とハイブリダイズする測定用核酸を含む測定用キットが提供される。

本発明によれば、本タンパク質に結合する単離抗体またはモノクローナル抗体が提供される。

【0016】

さらに、本発明によれば、生物試料における本発明のタンパク質または核酸を定量することを含む、生物試料の癌化を検出する方法を提供する。

【0017】

【発明の実施の形態】

以下、本発明の説明のために、好ましい実施形態に関して詳述する。

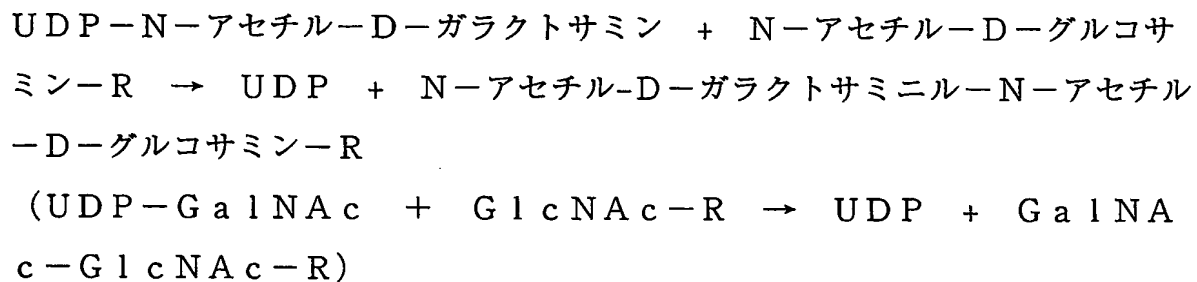
(1) タンパク質

下記実施例において詳述する方法によりクローニングされた、本発明のタンパク質をコードする核酸は、配列番号2または4に示される塩基配列を有し、それがコードする推定アミノ酸配列が、該塩基配列の下に記載されている。なお、配列番号1または3には、該アミノ酸配列のみが記載されている。

【0018】

下記実施例で得られた本発明のタンパク質（本明細書では「NGalNAc-T1」および「NGalNAc-T2」と命名）は、次の性質を有する酵素である。なお、本発明のタンパク質の性質およびその活性の測定方法は下記実施例において詳述されている。

作用： N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する。
触媒する反応を反応式で記載すると、



基質特異性： N-アセチルグルコサミン、例えばN-アセチルグルコサミン β 1-3-R（Rは、マンノースやp-ニトロフェノール等の水酸基とエーテル結合した残基）。

本発明によれば、N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する活性を有するタンパク質が提供される。本発明のタンパク質は、本明細書に記載した特徴を有する限り、その起源、製法等は限定されない。即ち、本発明のタンパク質は、天然産のタンパク質、遺伝子工学的手法により組換えDNAから発現されたタンパク質、あるいは化学合成タンパク質の何れでもよい。

【0019】

本発明のタンパク質は、典型的には、配列番号1に記載したアミノ酸残基1039個からなるアミノ酸配列、または配列番号3に記載したアミノ酸残基998個からなるアミノ酸配列を有する。しかしながら、天然のタンパク質の中には、それを生産する生物種の品種の違いや、生態型（ecotype）の違いによる遺伝子の変異、あるいはよく似たアイソザイムの存在等に起因して、1から複数個のアミノ酸変異を有する変異タンパク質が存在することは周知である。なお、

本明細書で使用する用語「変異タンパク質」とは、配列番号1または3に示されるアミノ酸配列において、1若しくは複数個のアミノ酸が置換し若しくは欠失し、または該アミノ酸配列に1若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有し、N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する活性を有するタンパク質等を意味する。ここで、「複数個」とは、好ましくは1-300個、より好ましくは1-100個、最も好ましくは1-50個である。一般的には、部位特異的な変異によってアミノ酸が置換された場合に、元々のタンパク質が有する活性は保存される程度に置換が可能なアミノ酸の個数は、好ましくは1-10個である。

【0020】

本発明のタンパク質は、クローニングされた核酸の塩基配列からの推定に基づいて、配列番号1若しくは3、並びに配列番号2若しくは4（下段）のアミノ酸配列を有するが、その配列を有するタンパク質のみに限定されるわけではなく、本明細書中に記載した特性を有する限り全ての相同タンパク質を含むことが意図される。同一性は、少なくとも50%以上、好ましくは60%以上、より好ましくは70%以上、さらに好ましくは80%以上、さらになお好ましくは90%以上、最も好ましくは95%以上である。

【0021】

本明細書において、同一性のパーセントは、例えば、Altschulら (Nuc l. Acids. Res. 25., p. 3389-3402, 1997) に記載されているBLASTプログラム、あるいはPearsonら (Proc. Nat l. Acad. Sci. USA, p. 2444-2448, 1988) に記載されているFASTAを用いて配列情報と比較し決定することが可能である。当該プログラムは、インターネット上でNational Center for Biotechnology Information (NCBI)、あるいはDNA Data Bank of Japan (DDBJ) のウェブサイトから利用することが可能である。各プログラムによる同一性検索の各種条件（パラメーター）は同サイトに詳しく記載されており、一部の設定を適宜変更することが可能であるが、検索は通常デフォルト値を用いて行う。なお、当業者に

用いられる、配列比較の他のプログラムもまた、使用可能である。

【0022】

一般的に、同様の性質を有するアミノ酸同士の置換（例えば、ある疎水性アミノ酸から別の疎水性アミノ酸への置換、ある親水性アミノ酸から別の親水性アミノ酸への置換、ある酸性アミノ酸から別の酸性アミノ酸への置換、あるいはある塩基性アミノ酸から別の塩基性アミノ酸への置換）を導入した場合、得られる変異タンパク質はもとのタンパク質と同様の性質を有することが多い。遺伝子組換え技術を使用して、このような所望の変異を有する組換えタンパク質を作製する手法は当業者に周知であり、このような変異タンパク質も本発明の範囲に含まれる。

【0023】

本発明のタンパク質は、例えば、後述の実施例に従って、本発明の核酸による配列番号2または4に記載のDNA配列を大腸菌、酵母、昆虫、または動物細胞に、それぞれの宿主で増幅可能な発現ベクターを用いて導入および発現させることにより、当該タンパク質を大量に得ることができる。

【0024】

本発明のタンパク質について、GENETYX（ゼネティックス社）による同一性検索を行うと、NGa1NAc-T1については、最も似ているCSGa1NAcT1とはC末端226アミノ酸で26.1%の同一性であり、一方、NGa1NAc-T2については、CSGa1NAc-T1とはC末端の431アミノ酸で21.6%、CSGa1NAc-T2とはC末端の224アミノ酸で25.0%の同一性である。したがって、本発明のタンパク質は新規なタンパク質であると考えられる。

【0025】

なお、GENETYXは、核酸解析、タンパク質解析用の遺伝情報処理ソフトウェアで、通常ホモロジー解析やマルチアラインメント解析の他、シグナルペプチド予測やプロモーター部位予測、二次構造予測が可能である。また、本明細書で用いたホモロジー解析プログラムは、高速・高感度な方法として多用されているLipman-Pearson法（Lipman, D. J. & Pearso

n, W. R. , Science, 277, 1435-1441 (1985)) を採用している。

【0026】

本発明によって、このタンパク質のアミノ酸配列およびそれをコードする DNA 配列が開示されれば、当該配列またはその一部を利用して、ハイブリダイゼーション、PCR 等の核酸増幅反応等の遺伝子工学的手法を用いて、他の生物種から同様の生理活性を有するタンパク質をコードする遺伝子を容易に単離することができる。このような場合、それらの遺伝子がコードする新規タンパク質も本発明の範囲に含まれる。

【0027】

なお、本発明のタンパク質は、そのアミノ酸配列が上述した通りのものであり、前記酵素活性を有するものであれば、タンパク質に糖鎖が結合していてもよい。

【0028】

より具体的には、後述する実施例 2 に記載したように、本発明のタンパク質における受容体基質の探索から、該タンパク質は、GlcNAc に GalNAc を転移する活性を有するものである。

【0029】

(2) 核酸

本発明の核酸は、一本鎖および二本鎖型両方の DNA、およびその RNA 相補体も含む。DNA には、例えば、天然由来の DNA、組換え DNA、化学合成した DNA、PCR によって増幅された DNA、およびそれらの組み合わせが含まれる。本発明の核酸としては、DNA が好ましい。

【0030】

本発明の核酸は、配列番号 1 または 3 に示されるアミノ酸をコードする核酸（その相補体を含む）である。典型的には、配列番号 2 または 4 の塩基配列（その相補体を含む）を有するが、これは本発明の一例を示すにすぎない下記の実施例で得られたクローンの塩基配列である。天然の核酸の中には、それを生産する生物種の品種の違いや、生態型の違いに起因する少数の変異やよく似たアイソザイ

ムの存在に起因する少数の変異が存在することは当業者に周知である。従って、本発明の核酸は、配列番号 2 または 4 に記載の塩基配列を有する核酸のみに限定されるわけではなく、本発明のタンパク質をコードする全ての核酸を包含する。

【0031】

特に、本発明によってこのタンパク質のアミノ酸配列およびそれをコードする DNA 配列が開示されれば、この配列またはその一部を利用して、ハイブリダイゼーションや核酸増幅反応等の遺伝子工学の基本的手法を用いて、他の生物種から同様の生理活性を有するタンパク質をコードする核酸を容易に単離することができる。このような場合、そのような核酸も本発明の範囲に含まれる。

【0032】

本明細書において、「ストリンジントな条件下」とは、中程度または高程度なストリンジントな条件においてハイブリダイズすることを意味する。具体的には、中程度のストリンジントな条件は、例えば、DNA の長さに基づき、一般の技術を有する当業者によって、容易に決定することが可能である。基本的な条件は、Sambrook ら, Molecular Cloning: A Laboratory Manual, 第 3 版, Vol. 1, 7.42-7.45 Cold Spring Harbor Laboratory Press, 2001 に示され、そしてニトロセルロースフィルターに関し、5XSSC、0.5% SDS、1.0mM EDTA (pH 8.0) の前洗浄溶液、約 40-50℃での、約 50%ホルムアミド、2XSSC-6XSSC (または約 42℃での約 50%ホルムアミド中の、スターク溶液 (Stark's solution) などの他の同様のハイブリダイゼーション溶液) のハイブリダイゼーション条件、および約 60℃、0.5XSSC、0.1% SDS の洗浄条件の使用が含まれる。高ストリンジントな条件もまた、例えば DNA の長さに基づき、当業者によって、容易に決定することが可能である。一般的に、こうした条件は、中程度にストリンジントな条件よりも高い温度および／または低い塩濃度でのハイブリダイゼーションおよび／または洗浄を含み、例えば上記のようなハイブリダイゼーション条件、およびおよそ 68℃、0.2XSSC、0.1% SDS の洗浄を伴うと定義される。当業者は、温度および洗浄溶液塩濃度は、プロ

ープの長さ等の要因に従って、必要に応じて調整可能であることを認識するであろう。

【0033】

核酸増幅反応は、例えば、ポリメラーゼ連鎖反応 (PCR) [Saiki R. K., et al., Science, 230, 1350-1354 (1985)]、ライゲース連鎖反応 (LCR) [Wu D. Y., et al., Genomics, 4, 560-569 (1989); Barringer K. J., et al., Gene, 89, 117-122 (1990); Barany F., Proc. Natl. Acad. Sci. USA, 88, 189-193 (1991)] および転写に基づく増幅 [Kwoh D. Y., et al., Proc. Natl. Acad. Sci. USA, 86, 1173-1177 (1989)] 等の温度循環を必要とする反応、並びに鎖置換反応 (SDA) [Walker G. T., et al., Proc. Natl. Acad. Sci. USA, 89, 392-396 (1992); Walker G. T., et al., Nuc. Acids. Res., 20, 1691-1696 (1992)]、自己保持配列複製 (3SR) [Guatelli J. C., Proc. Natl. Acad. Sci. USA, 87, 1874-1878 (1990)] および Q β レプリカーゼシステム [リザイルディら、BioTechnology 6, p. 1197-1202 (1988)] 等の恒温反応を含む。また、欧州特許第 0525882 号に記載されている標的核酸と変異配列の競合増幅による核酸配列に基づく増幅 (Nucleic Acid Sequence Based Amplification: NASABA) 反応等も利用可能である。好ましくは PCR 法である。

【0034】

上記のようなハイブリダイゼーション、核酸増幅反応等を使用してクローニングされる相同な核酸は、配列表の配列番号 2 または 4 に記載の塩基配列に対して少なくとも 50% 以上、好ましくは 60% 以上、より好ましくは 70% 以上、さらに好ましくは 80% 以上、さらになお好ましくは 90% 以上、最も好ましくは 95% 以上の同一性を有する。

【0035】

同一性パーセントは、視覚的検査および数学的計算によって決定することが可能である。あるいは、2つの核酸配列の同一性パーセントは、Devereuxら、Nuc l. A c i d s R e s. 12:387, 1984に記載され、そしてウイスコンシン大学遺伝学コンピューターグループ(UWGCG)より入手可能なGAPコンピュータープログラム、バージョン6.0を用いて、配列情報を比較することによって、決定可能である。GAPプログラムの好ましいデフォルトパラメーターには: (1) ヌクレオチドに関する単一(unary)比較マトリックス(同一に対し1および非同一直し0の値を含む)、およびSchwartzおよびDayhoff監修, Atlas of Protein Sequence and Structure, pp. 353-358, National Biomedical Research Foundation, 1979に記載されるような、GribnikovおよびBurgess, Nuc l. A c i d s R e s. 14:6745, 1986の加重比較マトリックス; (2) 各ギャップに対する3.0のペナルティおよび各ギャップ中の各記号に対しさらに0.10のペナルティ; および(3) 末端ギャップに対するペナルティなし、が含まれる。当業者に用いられる、配列比較の他のプログラムもまた、使用可能である。

【0036】

(3) 組換えベクターと形質転換体

本発明によればまた、本発明の核酸を含む組換えベクターが提供される。プラスミドなどのベクターに本発明の核酸のDNA断片を組み込む方法としては、例えば、Sambrook, J. ら, Molecular Cloning, A Laboratory Manual (3rd edition), Cold Spring Harbor Laboratory, 1.1 (2001)に記載の方法などが挙げられる。簡便には、市販のライゲーションキット(例えば、宝酒造製等)を用いることもできる。このようにして得られる組換えベクター(例えば、組換えプラスミド)は、宿主細胞(例えば、E-c o i l T B 1, L E 3 9 2 またはX L - 1 B l u e等)に導入される。

【0037】

プラスミドを宿主細胞に導入する方法としては、Sambrook, J. ら, Molecular Cloning, A Laboratory Manual (3rd edition), Cold Spring Harbor Laboratory, 16.1 (2001) に記載の塩化カルシウム法または塩化カルシウム／塩化ルビジウム法、エレクトロポレーション法、エレクトロインジェクション法、PEGなどの化学的な処理による方法、遺伝子銃などを用いる方法などが挙げられる。

【0038】

ベクターは、簡便には当業界において入手可能な組換え用ベクター（例えば、プラスミドDNAなど）に所望の遺伝子を常法により連結することによって調製することができる。用いられるベクターの具体例としては、大腸菌由来のプラスミドとして、例えば、pBluescript、pUC18、pUC19、pBR322などが例示されるがこれらに限定されない。

【0039】

所望のタンパク質を生産する目的においては、特に、発現ベクターが有用である。発現ベクターの種類は、原核細胞および／または真核細胞の各種の宿主細胞中で所望の遺伝子を発現し、所望のタンパク質を生産する機能を有するものであれば特に限定されないが、例えば、大腸菌用発現ベクターとして、pQE-30、pQE-60、pMAL-C2、pMAL-p2、pSE420などが好ましく、酵母用発現ベクターとしてpYES2（サッカロマイセス属）、pPIC3.5K、pPIC9K、pAO815（以上ピキア属）、昆虫用発現ベクターとしてpBacPAK8／9、pBK283、pVL1392、pBlueBac4.5などが好ましい。

【0040】

形質転換体は、所望の発現ベクターを宿主細胞に導入することにより調製することができる。用いられる宿主細胞としては、本発明の発現ベクターに適合し、形質転換され得るものであれば特に制限はなく、本発明の技術分野において通常使用される天然の細胞、または人工的に樹立された組換え細胞など種々の細胞を

用いることが可能である。例えば、細菌（エシェリキア属菌、バチルス属菌）、酵母（サッカロマイセス属、ピキア属など）、動物細胞、昆虫細胞、植物細胞などが挙げられる。

【0041】

宿主細胞は、大腸菌、酵母または昆虫細胞が好ましく、具体的には、大腸菌（M15、JM109、BL21等）、酵母（INVS c1（サッカロマイセス属）、GS115、KM71（以上ピキア属）など）、昆虫細胞（BmN4、カイコ幼虫など）などが例示される。また、動物細胞としてはマウス由来、アフリカツメガエル由来、ラット由来、ハムスター由来、サル由来またはヒト由来の細胞若しくはそれらの細胞から樹立した培養細胞株などが例示される。より具体的には、宿主細胞は、サル腎由来の細胞株であるCOS細胞が好ましい。

【0042】

宿主細胞として細菌、特に大腸菌を用いる場合、一般に発現ベクターは少なくとも、プロモーター／オペレーター領域、開始コドン、所望のタンパク質をコードする遺伝子、終止コドン、ターミネーターおよび複製可能単位から構成される。

【0043】

宿主細胞として酵母、植物細胞、動物細胞または昆虫細胞を用いる場合には、一般に発現ベクターは少なくとも、プロモーター、開始コドン、所望のタンパク質をコードする遺伝子、終止コドン、ターミネーターを含んでいることが好ましい。またシグナルペプチドをコードするDNA、エンハンサー配列、所望の遺伝子の5'側および3'側の非翻訳領域、選択マーカー領域または複製可能単位などを適宜含んでもよい。

【0044】

本発明のベクターにおいて、好適な開始コドンとしては、メチオニンコドン（ATG）が例示される。また、終止コドンとしては、常用の終止コドン（例えば、TAG、TGA、TAAなど）が例示される。

【0045】

複製可能単位とは、宿主細胞中でその全DNA配列を複製することができる能

力をもつDNAを意味し、天然のプラスミド、人工的に修飾されたプラスミド（天然のプラスミドから調製されたプラスミド）および合成プラスミド等が含まれる。好適なプラスミドとしては、*E. coli*ではプラスミドpQE30、pETまたはpCALもしくはそれらの人工的修飾物（pQE30、pETまたはpCALを適当な制限酵素で処理して得られるDNAフラグメント）が、酵母ではプラスミドpYES2もしくはpPIC9Kが、また昆虫細胞ではプラスミドpBacPAK8/9等があげられる。

【0046】

エンハンサー配列、ターミネーター配列については、例えば、それぞれSV40に由来するもの等、当業者において通常使用されるものを用いることができる。

【0047】

選択マーカーとしては、通常使用されるものを常法により用いることができる。例えばテトラサイクリン、アンピシリン、またはカナマイシンもしくはネオマイシン、ハイグロマイシンまたはスペクチノマイシン等の抗生物質耐性遺伝子などが例示される。

【0048】

発現ベクターは、少なくとも、上述のプロモーター、開始コドン、所望のタンパク質をコードする遺伝子、終止コドン、およびターミネーター領域を連続的かつ環状に適当な複製可能単位に連結することによって調製することができる。またこの際、所望により制限酵素での消化やT4DNAリガーゼを用いるライゲーション等の常法により適当なDNAフラグメント（例えば、リンカー、他の制限酵素部位など）を用いることができる。

【0049】

本発明の発現ベクターの宿主細胞への導入〔形質転換（形質移入）〕は従来公知の方法を用いて行うことができる。

例えば、細菌（*E. coli*, *Bacillus subtilis*等）の場合は、例えばCohenらの方法〔*Proc. Natl. Acad. Sci. USA*, 69, 2110（1972）〕、プロトプラスト法〔*Mol. Gen.*

Genet., 168, 111 (1979)] やコンピテント法 [J. Mol. Biol., 56, 209 (1971)] によって、Saccharomyces cerevisiaeの場合は、例えばHinnenらの方法 [Proc. Natl. Acad. Sci. USA, 75, 1927 (1978)] やリチウム法 [J. Bacteriol., 153, 163 (1983)] によって、植物細胞の場合は、例えばリーフディスク法 [Science, 227, 129 (1985)]、エレクトロポレーション法 [Nature, 319, 791 (1986)] によって、動物細胞の場合は、例えばGrahamの方法 [Virology, 52, 456 (1973)]、昆虫細胞の場合は、例えばSummersらの方法 [Mol. Cell. Biol., 3, 2156-2165 (1983)] によってそれぞれ形質転換することができる。

【0050】

(4) タンパク質の単離・精製

本発明のタンパク質は、上記の如く調整された発現ベクターを含む形質転換細胞を栄養培地で培養することによって発現（生産）することができる。栄養培地は、宿主細胞（形質転換体）の生育に必要な炭素源、無機窒素源もしくは有機窒素源を含んでいることが好ましい。炭素源としては、たとえばグルコース、デキストラン、可溶性デンプン、ショ糖、メタノールなどが、例示される。無機窒素源もしくは有機窒素源としては、例えばアンモニウム塩類、硝酸塩類、アミノ酸、コーンステープ・リカー、ペプトン、カゼイン、肉エキス、大豆粕、バレイショ抽出液などが例示される。また、所望により他の栄養素（例えば無機塩（例えば、塩化ナトリウム、塩化カルシウム、リン酸二水素ナトリウム、塩化マグネシウム）、ビタミン類、抗生物質（例えばテトラサイクリン、ネオマイシン、アンピシリン、カナマイシン等）など）を含んでいてもよい。培養は、当業界において知られている方法により行われる。培養条件、例えば温度、培地のpH及び培養時間は、本発明のタンパク質が大量に生産されるように適宜選択される。

【0051】

本発明のタンパク質は、上記培養により得られる培養物より以下のようにして取得することができる。すなわち、本発明のタンパク質が宿主細胞内に蓄積する

場合には、遠心分離やろ過などの操作により宿主細胞を集め、これを適当な緩衝液（例えば濃度が10M～100mM程度のトリス緩衝液、リン酸緩衝液、HEPES緩衝液、MES緩衝液などの緩衝液。pHは用いる緩衝液によって異なるが、pH5.0～9.0の範囲が望ましい）に懸濁した後、用いる宿主細胞に適した方法で細胞を破壊し、遠心分離により宿主細胞の内容物を得る。一方、本発明のタンパク質が宿主細胞外に分泌される場合には、遠心分離やろ過などの操作により宿主細胞と培地を分離し、培養液を得る。宿主細胞破壊液、あるいは培養液はそのまま、または硫酸沈殿と透析を行なった後に、本発明のタンパク質の単離・精製に供することができる。単離・精製の方法としては、以下の方法が挙げることができる。即ち、当該タンパクに6×ヒスチジンやGST、マルトース結合タンパクといったタグを付けている場合には、一般に用いられるそれぞれのタグに適したアフィニティークロマトグラフィーによる方法を挙げることができる。一方、そのようなタグを付けずに本発明のタンパク質を生産した場合には、例えば後述する実施例に詳しく述べられている方法、即ちイオン交換クロマトグラフィーによる方法を挙げることができる。また、これに加えてゲルろ過や疎水性クロマトグラフィー、等電点クロマトグラフィーなどを組み合わせる方法も挙げることができる。

【0052】

本発明のタンパク質を、N-アセチルグルコサミンを有する糖タンパク質、オリゴ糖または多糖等に作用させることにより、N-アセチルガラクトサミンが転移される。従って、本発明のタンパク質は、糖タンパク質の糖鎖の修飾や、糖類の合成に用いることができる。さらに、このタンパク質を免疫原として動物に投与することにより、該タンパク質に対する抗体を作製することができ、該抗体を用いて免疫測定法により該タンパク質を測定することが可能になる。従って、本発明のタンパク質およびこれをコードする核酸は、このような免疫原の作製に有用である。

【0053】

本発明のタンパク質はまた、精製および同定を容易にするために添加されるペプチドを含んでもよい。こうしたペプチドには、例えば、ポリ-Hisまたは米

国特許第5, 011, 912号およびHoppら, Bio/Technology, 6:1204, 1988に記載される抗原性同定ペプチドが含まれる。こうしたペプチドの1つはFLAG (登録商標) ペプチド、Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lysであり、該ペプチドは非常に抗原性であり、そして特異的なモノクローナル抗体が可逆的に結合するエピトープを提供し、発現された組換えタンパク質の迅速なアッセイおよび容易な精製を可能にする。4E11と称されるネズミハイブリドーマは、本明細書に援用される米国特許第5, 011, 912号に記載されるように、特定の二価金属陽イオンの存在下で、FLAG (登録商標) ペプチドに結合するモノクローナル抗体を産生する。4E11ハイブリドーマ細胞株は、寄託番号HB 9259下に、アメリカン・タイプ・カルチャー・コレクション (American Type Culture Collection) に寄託されている。FLAG (登録商標) ペプチドに結合するモノクローナル抗体は、Eastman Kodak Co., Scientific Imaging Systems Division、コネチカット州ニューヘブーンより入手可能である。

【0054】

具体的には、後述される実施例1に記載されるように、本発明のタンパク質を発現する発現ベクターにFLAGのcDNAを挿入し、FLAG標識したタンパク質を発現させ、抗FLAG抗体によって、該タンパク質の発現を確認することができる。

【0055】

(5) 測定用核酸

本発明によれば、本発明の核酸とハイブリダイズする核酸 (以下、「測定用核酸」と称する) が提供される。本発明の測定用核酸は、典型的には、本発明のタンパク質をコードする核酸の天然由来のまたは合成されたフラグメントであり、プライマーまたはプローブを含むが、これらに限定されるものではない。なお、本明細書において使用される用語「測定」には、検出、増幅、定量、および半定量のいずれもが包含される。

【0056】

(a) プライマー

本発明の測定用核酸を核酸増幅反応用のプライマーとして使用する場合、本発明の測定用核酸は、オリゴヌクレオチドであって、

配列番号 1 または 3 に示すタンパク質をコードする遺伝子の塩基配列から以下の条件を満たすように 2 つの領域を選択し：

- 1) 各領域の長さが 15 - 50 塩基であること；
- 2) 各領域中の G + C の割合が 40 - 70 % であること；

上記領域と同じ塩基配列若しくは上記領域に相補的な塩基配列を有する一本鎖 DNA を製造し、または、上記一本鎖 DNA によってコードされるアミノ酸残基を変化させないように遺伝子暗号の縮重を考慮した一本鎖 DNA の混合物を製造し、さらに必要であれば上記タンパク質をコードする遺伝子の塩基配列に対する結合特異性を失わないように修飾した上記一本鎖 DNA を製造することを含む方法により製造された当該オリゴヌクレオチドが提供される。

【0057】

本発明のプライマーは、本発明の核酸の部分領域と相同的な配列を有することが好ましいが、1 または 2 塩基の不一致があっても差し支えない。

なお、本発明のプライマーの塩基数は 15 塩基以上、好ましくは 18 塩基以上、より好ましくは 21 塩基以上であり、50 塩基以下である。

【0058】

本発明のプライマーは、典型的には、配列番号 20、21、23 および 24 からなる群から選択された塩基配列を有し、単独、または適当な 2 種を組み合わせたプライマー対として用いることができる。このヌクレオチド配列は、配列番号 1 または 3 のアミノ酸配列に基づいて、各々のタンパク質をコードする遺伝子断片のクローニングのための PCR 用プライマーとして設計したものであり、当該アミノ酸をコードすることが可能な全ての塩基をミックスしたプライマーである。

【0059】

(b) プローブ

本発明の測定用核酸をプローブとして使用する場合、本発明の測定用核酸は、配列番号 2 若しくは 4 に記載の塩基配列の全体または部分領域と相同的な配列を

有することが好ましいが、1または2塩基の不一致があっても差し支えない。なお、本発明のプロープの塩基数は15塩基以上、好ましくは20塩基以上であり、コード領域の全長、即ち、3120塩基（配列番号2に対応）または2997塩基（配列番号4に対応）以下である。本発明のプライマーは、典型的には、配列番号22または24に記載された塩基配列を有する。プロープは天然由来の核酸より、制限酵素処理によって得てもよいし、合成したオリゴヌクレオチドであってもよい。

【0060】

本発明のプロープには、該プロープが標的配列とハイブリダイズしたことを検出または確認するために、蛍光標識、放射標識、ビオチン標識等の標識を付した標識プロープが含まれる。被検核酸またはその増幅物を固相化し、標識プロープとハイブリダイズさせ、洗浄後、固相に結合された標識を測定することにより、検体中に被検核酸が存在するかを決定することができる。あるいは、測定用核酸を固相化し、被検核酸をハイブリダイズさせ、固相に結合した被検核酸を標識プロープ等で検出することも可能である。このような場合、固相に結合した測定用核酸もプロープと呼ぶ。

【0061】

一般的に、PCRのような核酸増幅法自体は、この分野において周知であり、そのための試薬キットおよび装置も市販されているので容易に行うことができる。上記した本発明の測定用核酸の一对をプライマーとして用い、被検核酸を鋳型として用いて核酸増幅法を行うと、被検核酸が増幅されるのに対し、検体中に被検核酸が含まれない場合には増幅が起きないので、増幅産物を検出することにより検体中に被検核酸が存在するか否かを知ることができる。増幅産物の検出は、増幅後の反応溶液を電気泳動し、バンドをエチジウムブロミド等で染色する方法や、電気泳動後の増幅産物をナイロン膜等の固相に不動化し、被検核酸と特異的にハイブリダイズする標識プロープとハイブリダイズさせ、洗浄後、該標識を検出することにより行うことができる。また、クエンチャー蛍光色素とレポーター蛍光色素を用いたいわゆるリアルタイム検出PCRを行うことにより、検体中の被検核酸の量を定量することも可能である。なお、リアルタイム検出PCR用の

キットも市販されているので、容易に行うことができる。さらに、電気泳動バンドの強度に基づいて被検核酸を半定量することも可能である。なお、被検核酸は、mRNAでも、mRNAから逆転写したcDNAであつてもよい。被検核酸としてmRNAを増幅する場合には、上記一对のプライマーを用いたNASBA法(3SR法、TMA法)を採用することもできる。NASBA法自体は周知であり、そのためのキットも市販されているので、上記一对のプライマーを用いて容易に実施することができる。

【0062】

(6) 抗体

本発明のタンパク質に免疫反応性である抗体が本明細書に提供される。こうした抗体は、(非特異的結合と対照的に)抗体の抗原結合部位を介して、該タンパク質に特異的に結合する。したがって、上述のような、配列番号1および配列番号3のタンパク質、断片、変異体、融合タンパク質などを、それと免疫反応性である抗体を産生する際の「免疫原」として使用することが可能である。より具体的には、タンパク質、断片、変異体、融合タンパク質などは、抗体形成を引き出す抗原決定基またはエピトープを含む。これらの抗原決定基またはエピトープは、直鎖でも高次構造的(conformational)(断続的)でもどちらでもよい。なお、該抗原決定基またはエピトープは、当該技術分野に知られるいかなる方法によって同定してもよい。

【0063】

したがって、本発明の1つの側面は、本発明のタンパク質の抗原性エピトープに関する。こうしたエピトープは、以下により詳細に記載されるように、抗体、特にモノクローナル抗体を作成するのに有用である。さらに、本発明のタンパク質由来のエピトープは、アッセイにおいて、そしてポリクローナル血清または培養ハイブリドーマ由来の上清などの物質から特異的に結合する抗体を精製する研究試薬として使用可能である。こうしたエピトープまたはその変異体は、固相合成、タンパク質の化学的または酵素的切断などの、当該技術分野に公知の技術を用いて、あるいは組換えDNA技術を用いて、産生することが可能である。

【0064】

本発明のタンパク質によって誘導される可能性がある抗体に関しては、該タンパク質の全部若しくは一部が単離されていても、またはエピトープが単離されていても、ポリクローナル抗体およびモノクローナル抗体はどちらも、慣用的技術によって調製することが可能である。例えば、Kennetら（監修），*Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, ニューヨーク, 1980を参照されたい。

【0065】

本発明のタンパク質に特異的なモノクローナル抗体を産生するハイブリドーマ細胞株もまた、本明細書に意図される。こうしたハイブリドーマは、慣用的技術によって産生しそして同定することが可能である。こうしたハイブリドーマ細胞株を産生するための1つの方法は、動物を該タンパク質で免疫し；免疫された動物から脾臓細胞を採取し；前記脾臓細胞を骨髓腫細胞株に融合させ、それによりハイブリドーマ細胞を生成し；そして該タンパク質に結合するモノクローナル抗体を産生するハイブリドーマ細胞株を同定することを含む。モノクローナル抗体は、慣用的技術によって回収可能である。

【0066】

本発明のモノクローナル抗体には、キメラ抗体、例えば、ネズミモノクローナル抗体のヒト化型が含まれる。こうしたヒト化抗体を既知の技術によって調製し、そして抗体がヒトに投与されるとき、免疫原性の減少という利点を提供してもよい。1つの態様において、ヒト化モノクローナル抗体は、ネズミ抗体の可変領域（またはその抗原結合部位のみ）およびヒト抗体由来の定常領域を含む。あるいは、ヒト化抗体断片は、ネズミモノクローナル抗体の抗原結合部位およびヒト抗体由来の可変領域断片（抗原結合部位を欠く）を含んでもよい。

【0067】

慣用的技術によって産生可能な、抗体の抗原結合断片もまた、本発明に含まれる。こうした断片の例には、限定されるわけではないが、FabおよびF(ab')₂断片が含まれる。遺伝子工学技術によって産生される抗体断片および誘導体もまた提供される。

【0068】

1つの態様において、抗体は本発明のタンパク質に特異的であり、そして他のタンパク質と交差反応しない。こうした抗体を同定することが可能なスクリーニング法が公知であり、そして例えば、免疫アフィニティークロマトグラフィーを伴ってもよい。

【0069】

本発明の抗体は、*in vitro*または*in vivo*いずれかで、本発明のタンパク質または断片の存在を検出するアッセイで用いることが可能である。抗体はまた、免疫アフィニティークロマトグラフィーによって、本発明のポリペプチドまたは断片を精製する際にも使用可能である。

【0070】

さらに、結合パートナー、例えば受容体基質への本発明のタンパク質の結合を遮断することが可能な抗体を用いて、こうした結合から生じる生物学的活性を阻害することが可能である。こうした遮断抗体は、受容体基質を発現している特定の細胞への該タンパク質の結合を阻害する能力に関して、抗体を試験することによるなど、いかなる適切なアッセイ法を用いて、同定してもよい。あるいは、遮断抗体は、標的細胞の結合パートナーに結合している本発明のタンパク質から生じる生物学的影響を阻害する能力に関するアッセイにおいて、同定することが可能である。

【0071】

こうした抗体を、*in vitro*法で使用するか、または*in vivo*で投与して、抗体を生成した実体によって仲介される生物学的活性を阻害することが可能である。したがって、本発明のタンパク質と結合パートナーとの相互作用によって、（直接または間接的に）引き起こされるかまたは悪化される障害を治療することが可能である。療法は、結合パートナー仲介生物学的活性を阻害するのに有効な量の遮断抗体を、哺乳動物に*in vivo*投与することを伴う。一般的に、こうした療法の使用には、モノクローナル抗体が好ましい。1つの態様において、抗原結合抗体断片が使用される。

【0072】

(6) 癌マーカーおよび検出方法

本発明のタンパク質または核酸は、癌マーカーとして使用することができ、癌の診断、治療等に応用することが可能である。ここで、本明細書において使用される用語「癌」には、典型的には、悪性腫瘍全般をいい、該悪性腫瘍による疾病状態を含む。限定されるわけではないが、「癌」としては、例えば肺癌、肝臓癌、腎臓癌、白血病が含まれる。

【0073】

本明細書において「癌マーカー」というときは、生物試料が癌化されている場合に、癌化されていない生物試料と比べて多く発現される本発明のタンパク質または核酸をいう。なお、「生物試料」というときには、組織、器官および細胞を含む。好ましくは血液、より好ましくは病理組織である。

【0074】

具体的には、本発明のタンパク質を癌マーカーとする場合、本発明の検出方法には、(a) 生物試料の前記タンパク質を定量し；そして (b) 生物試料中の前記タンパク質の定量値が、対照の生物試料中の前記タンパク質の定量値よりも大きい場合には癌化していると判断する工程が含まれる。前記検出方法には、生物試料の該タンパク質を定量するために、本発明の抗体を使用することができる。本発明によれば、一般的には、タンパク質を定量する方法は、上記方法に限定されるものではなく、当該技術分野において周知の定量方法、例えば、ELISA法、ウェスタンブロットを使用することが可能である。前記定量値の比は、好ましくは1.5倍以上、より好ましくは3倍以上、さらに好ましくは10倍以上である。

【0075】

一方、本発明の核酸を癌マーカーとする場合、本発明の検出方法は、(a) 生物試料中の前記核酸を定量し、そして (b) 生物試料中の前記核酸の定量値が、対照の生物試料中の前記核酸の定量値の1.5倍以上である場合には癌化していると判断する工程を含む。好ましくは、(a) 生物試料中の前記核酸と少なくとも1つの前記測定用核酸とをハイブリダイズさせ；(b) 前記核酸を増幅させ；(c) 増幅産物に前記測定用核酸をハイブリダイズさせ；(d) 前記増幅産物と

ハイブリダイズした前記測定用核酸から生じるシグナルを定量し；そして（e）前記シグナルの定量値が、対照の試料生物中の対応するシグナルの定量値1.5倍以上である場合には癌化していると判断する工程を含む。

【0076】

より具体的には、後述する実施例4に記載されているように、癌化されている組織と正常な組織での核酸の発現レベルの比を定量PCRによって測定することによって癌化していることを判断することができる。本発明によれば、核酸の定量には、これに限定されるものではなく、例えば、RT-PCR、ノーザンブロット、ドットブロット、DNAマイクロアレイを使用することができる。定量には、同一組織等に広く一般的に存在する遺伝子の核酸、例えばグリセルアルデヒド-3リン酸-脱水素酵素（GAPDH）、 β -アクチンをコードする核酸を対照として利用することができる。癌化していると判断されるシグナルの定量比は、1.5以上であり、好ましくは3以上、さらにより好ましくは10以上である。

【0077】

以下、本発明を実施例に基づきより具体的に説明する。もっとも、本発明は下記実施例に限定されるものではない。

【0078】

【実施例】

実施例1

本発明のタンパク質の調製

1. 遺伝子データベースの検索と新規N-アセチルガラクトサミン転移酵素遺伝子の塩基配列決定

既存の β -1, 4-ガラクトース転移酵素遺伝子を用いて、遺伝子データベースから類似遺伝子の検索を行った。用いた配列は β -1, 4-ガラクトース転移酵素遺伝子の配列番号：AL161445、AF038660、AF038661、AF022367、AF038663、AF038664である。また検索は、Blast [Altschul et al., J. Mol. Biol., 215, 403-410 (1990)]等のプログラムを利用した。

【0079】

その結果、EST配列GenBank Accetion No. N48738およびゲノム配列GenBank Accetion No. AC006205が見出された。その後の検討の結果、両配列は異なる遺伝子を含んでいると考えられた（本明細書ではN48738を含む遺伝子をNGa1NAc-T1、AC006205に含まれる遺伝子をNGa1NAc-T2と呼ぶ）。両遺伝子とも翻訳開始点が不明で、全長を予測することが不可能であったので、CLONT ECH社のMarathon-Ready cDNA (Human Brain 若しくはStomach) を用いてコード領域情報の取得（5' RACE: Rapid Amplification of cDNA Ends）並びにクローニングを行った。

【0080】

NGa1NAc-T1のコード領域情報の取得

Marathon cDNA付属のAP1プライマーと（cDNA断片の両側にAP1、AP2のアダプターがついている）、見出した配列部分に設定したプライマーK12R6（5' -GCT CCT GCA GCT CCA GCT CCA-3'（配列番号5））でPCR（94℃20秒、60℃30秒、72℃2分を30サイクル）を行った。さらに、Marathon cDNA付属のAP2プライマーと、配列部分に設定したプライマーK12R5（5' -AAG CGA CTC CCT CGC GCC GAG T-3'（配列番号6））でnested PCR（94℃20秒、60℃30秒、72℃2分を30サイクル）を行った。その結果得られた約0.6 kbの断片を常法により精製し、塩基配列を解析した。しかし、糖転移酵素特有の膜貫通配列（疎水的な20アミノ酸）が出現しなかったため、得られた配列と後述のNGa1NAc-T2の塩基配列を参考にゲノムデータベース検索でEST配列（GenBank Accession No. PF058197）を見出した。この塩基配列情報をもとに2種のプライマー（K12F101: 5' -ATG CCG CGG CTC CCG GTG AAG AAG-3'（配列番号7）とK12R5）によりRT-PCRを行い、増幅が確認できたことから、このEST配列と5' RACEで得られた配列は1本のmRNA上に存在すること明らかになった。全長3120塩基のヌクレオチド配列を配列番

号2に示した。

【0081】

NGalNAc-T2のコード領域情報の取得

Marathon cDNA付属のAP1プライマーと(cDNA断片の両側にAP1、AP2のアダプターがついている)、見出した配列部分に設定したプライマーK13-R3 (5'-CAA CAG TTC AAG CTC CAG GAG GTA-3' (配列番号8))でPCR (94℃20秒、60℃30秒、72℃2分を30サイクル)を行った。さらに、Marathon cDNA付属のAP2プライマーと、配列部分に設定したプライマーK13R2 (5'-CTG ACG CTT TTC CAC GTT CAC AAT-3' (配列番号9)) nested PCR (94℃20秒、60℃30秒、72℃2分を30サイクル)を行なった。その結果得られた約1.0 kbの断片を常法により精製し、塩基配列を解析し、タンパク質のコード領域を決定した。しかし、糖転移酵素特有の膜貫通配列(疎水的な20アミノ酸)が出現しなかったため、さらに3度の5' RACEを行った。ここで使用したプライマーを表2に示した。

【0082】

その結果、得られた全長2997塩基のヌクレオチド配列を配列番号4に示した。

【0083】

【表2】

表2 RACEに使用した各種プライマー

2度目の5' RACEプライマー

K13 R6 5' -CAC CCC GTC TCT GCT CTG CGA T-3' (配列番号10)

K13 R5 5' -GTC TTC CTG GGG CTG TCA CCA-3' (配列番号11)

3度目の5' RACEプライマー

K13 R7 5' -CAC CTC ATC CAT CTG TAG GAA CGT-3' (配列番号12)

K13 R8 5' -CTG TCG CCA TGC AAC TTC CAC GT-3' (配列番号13)

4度目の5' RACEプライマー

K13 R12 5' -AAT GTC GTG GTC CTC GAG GCT CA-3' (配列番号14)

K13 R11 5' -GAT GGT AGA ACT GGA GGT GTG GAT-3' (配列番号15)

2. GalNAc-T遺伝子の発現ベクターへの組込み

GalNAc-Tの発現系を作成するため、まずGalNAc-T遺伝子の一部をシグマ社のpFLAG-CMV1に組込んだ。以下に詳細を述べる。

【0084】

NGalNAc-T1のpFLAG-CMV1への組込み

Marathon cDNA (Human Brain) を鋳型として、フォワードプライマーK12-Hin-F2: 5' -CCC AAG CTT CGG GGG GTC CAC GCT GCG CCA T-3' (配列番号16)、リバープライマーK12-Xba-R1: 5' -GCT CTA GAC TCA AGA CGC CCC CGT GCG AGA-3' (配列番号17) を用いて、宝酒造LA Taq DNAポリメラーゼにより配列番号1若しくは2の62番目から1039番目のアミノ酸に相当する領域を増幅した。この断片を、プライマー中に含まれる制限部位(HindIIIとXbaI)で切断し、HindIIIとXbaIで切断したpFLAG-CMV1に東洋紡社のLigation Highを用いて挿入し、pFLAG-NGalNAc-T1を作製した。

【0085】

GalNAc-T2のpFLAG-CMV1への組込み

Marathon cDNA (Human Stomach) を鋳型として、フォワードプライマーK13-Eco-F1: 5' -GGA ATT CGA GGT ACG GCA GCT GGA GAG AA-3' (配列番号18)、リバープライマーK13-Sal-R1: 5' -ACG CGT CGA CCT ACA GCG TCT TCA TCT GGC GA-3' (配列番号19) を用いて、宝酒造LA Taq DNAポリメラーゼにより配列番号3若しくは4の57番目から998番目のアミノ酸に相当する領域を増幅した。この断片を、プライマー中に含まれる制限部位(EcoRIとSalI)で切断し、EcoRIとSalIで切断したpcDNA3.1に一旦挿入した。これをEcoRIとPmeIで切断し、NGalNAc-T2の活性部位を含む断片をpF

LAG-CMV1のEcoRI-EcoRV部位に東洋紡社のLigation Highを用いて挿入し、pFLAG-NGalNAc-T2を作製した。

【0086】

3. トランスフェクションとリコンビナント酵素の発現

15 μ gのpFLAG-NGalNAc-T1若しくはpFLAG-NGalNAc-T2を、 2×10^6 のCOS-1細胞を10% FCS (ウシ胎児血清)を含むDMEM (Dulbecco's modified Eagle's medium)で1晩培養したものにリポフェクタミン2000 (インビトロジェン社)を用い、同社のプロトコルに従い導入した。48時間から72時間培養した上清を回収し、上清10 mlにNaN₃ (0.05%)、NaCl (150 mM)、CaCl₂ (2 mM)、抗M1レジン (シグマ社) (50 μ l)を混合し、4℃で一夜攪拌した。反応混合液を遠心分離して (3000 rpm、5分、4℃) ペレットを回収し、2 mMのCaCl₂・TBSを900 μ l加えて再度遠心分離 (2000 rpm、5分、4℃) し、ペレットを200 μ lの1 mM CaCl₂・TBSに浮遊させ活性測定のサンプル (NGalNAc-T1若しくはNGalNAc-T2酵素液)とした。

【0087】

酵素液は常法により、SDS-PAGEとウエスタンブロッティングを行い、目的とするタンパク質の発現を確認した。抗体は抗FLAG M2-ペルオキシダーゼ (A-8592、SIGMA社)を用いた。

【0088】

実施例2

本発明酵素を用いた活性測定

1. 供与体基質の探索

各種単糖受容体基質の混合液に対し、酵素液5 μ lと様々な供与体基質を用いて、本発明酵素の供与体基質の探索を行った。

【0089】

受容体基質は、Gal- α -pNp、Gal- β -oNp、GalNAc- α -Bz、GalNAc- β -pNp、GlcNAc- α -pNp、GlcNAc

α -pNp、Glc- α -pNp、Glc- β -pNp、GlcA- β -pNp、Fuc- α -pNp、Man- α -pNp（以上、CALBIOCHEM社）、Xyl- α -pNp、Xyl- β -pNp（以上、シグマ社）が各2.5 nmol/20 μ l 含まれるように調製した。また、各種供与体基質（UDP-GalNAc、UDP-GlcNAc、UDP-Gal、GDP-Man、UDP-GlcA、UDP-Xyl および GDP-Fuc、何れもシグマ社）に対する反応液は表3にまとめたとおりである。

【0090】

【表3】

GalNAc-T	
MES or HEPES (pH 5.5 ~)	50 mM
UDP-GalNAc	0.5 mM
UDP-[14C]GalNAc	2 nCi/ul
MnCl ₂	20 mM
Triton X-100	0.5%

GlcNAc-T	
HEPES (pH 7.0 or 7.5)	14 mM
UDP-GlcNAc	0.5 mM
UDP-[14C]GlcNAc	2 nCi/ul
MnCl ₂	10 mM
Triton CF-54	0.5%
ATP	0.75 mM

Gal-T	
HEPES (pH 7.0 or 7.5)	14 mM
UDP-Gal	0.25 mM
UDP-[14C]Gal	2.5 nCi/ul
MnCl ₂	10 mM
ATP	0.75 mM

GlcA-T	
MES (pH 7.0)	50 mM
UDP-GlcA	0.25 mM
UDP-[14C]GlcA	2 nCi/ul
MnCl ₂	10 mM

Xyl-T	
MES (pH 7.0)	50 mM
UDP-Xyl	0.25 mM
UDP-[14C]Xyl	1 nCi/ul
MnCl ₂	10 mM

Fuc-T	
cacodylate buffer (pH 7.0)	50 mM
GDP-[14C]Fuc	1 nCi/ul
MnCl ₂	10 mM
ATP	5 mM

Man-T	
Tris (pH 7.2)	50 mM
GDP-[14C]Man	2 nCi/ul
MnCl ₂	10 mM
Triton X-100	0.6%

【0091】

また、反応時間はすべて16時間とした。反応後、Sep Pack C18 カラム（ウォーターズ社）で未反応の放射活性を有する供与体基質を除去し、受容器に取り込まれた供与体由来の放射活性を液体シンチレーションカウンターで測定した。その結果、NGalNAc-T1、NGalNAc-T2とも、UDP-GlcAでも若干のバックグラウンドが見られたが、UDP-GalNAcを供与体基質に用いたときのみ、最も高い活性が検出できた。

【0092】

2. 受容体基質の探索

さらに、受容体基質を調べる目的で、それぞれの受容体基質を単独で使用し（ $10 \text{ nmol} / 20 \mu\text{l}$ ）、反応を行った。その結果、NGalNAc-T1の場合に256.26 dpm、NGalNAc-T2の場合に1221.22 dpmと、GlcNAc- β -pNpを用いたときのみ、有意な放射活性が検出された。以上の結果から、NGalNAc-T1およびNGalNAc-T2は共にGlcNAc-TにGalNAcを転移する糖転移酵素であることが明らかになった。

【0093】

3. 至適pHの検討

以上のように、NGalNAc-T1およびNGalNAc-T2がGlcNAcにGalNAcを転移する糖転移酵素であることが明らかになった。そこで、両酵素の至適pHを検討した。ここで用いた緩衝液はMES（pH 5.5、6.0、6.26、6.5、6.75）、HEPES（pH 6.75、7.0、7.4）である。その結果、表4にまとめたように、pH 6.5で活性が高くなる傾向がみられた。

表4 NGalNAc-T1およびNGalNAc-T2の酵素活性における至適pHの検討結果

【0094】

【表 4】

NGalNAc-T1

pH	放射活性取込み量 (A)	ブランク (B)	(A) - (B)
MES緩衝液 (pH 5.5)	339.76	263.21	76.55
MES緩衝液 (pH 6.0)	321.04	263.21	57.83
MES緩衝液 (pH 6.26)	636.34	263.21	373.13
MES緩衝液 (pH 6.5)	1767.72	263.21	1504.51
MES緩衝液 (pH 6.75)	923.92	263.21	660.71
HEPES緩衝液 (pH 6.75)	1685.06	263.21	1421.85
HEPES緩衝液 (pH 7.0)	1138.38	263.21	875.17
HEPES緩衝液 (pH 7.4)	2587.48	263.21	2324.27

(dpm)

NGalNAc-T2

pH	放射活性取込み量 (A)	ブランク (B)	(A) - (B)
MES緩衝液 (pH 5.5)	336.20	263.21	72.99
MES緩衝液 (pH 6.0)	341.92	263.21	78.71
MES緩衝液 (pH 6.26)	339.50	263.21	76.29
MES緩衝液 (pH 6.5)	753.26	263.21	490.05
MES緩衝液 (pH 6.75)	529.24	263.21	266.03
HEPES緩衝液 (pH 6.75)	915.16	263.21	651.95
HEPES緩衝液 (pH 7.0)	786.70	263.21	523.49
HEPES緩衝液 (pH 7.4)	586.32	263.21	323.11

(dpm)

【0095】

なお、ここでは、酵素を添加しない場合のブランク値に、MES (pH 6.75) のときの値 (263.21 dpm) を採用した。また、NGalNAc-T1でpH 7.4のときに最も高い値を示したが、必ずしも高いpHにすると活性も高くなるわけではないので、今後の実験にはMES (pH 6.5) を使用することにした。

【0096】

4. 2価陽イオン要求性の検討

一般に、糖転移酵素には2価の陽イオンを要求するものが多い。そこで、各種2価陽イオンを添加して、それぞれの酵素の活性を検討した。その結果、NGalNAc-T1では Mn^{2+} 、NGalNAc-T2では Mg^{2+} 、 Mn^{2+} 、 Co^{2+} を添加したときに高い活性を示した (表5参照)。これに対し、キレート剤であるEDTAを添加するとどちらの酵素も活性を示さなかった。以上の結果から両酵素が2価陽イオンを要求することが明らかになった。

表5 NGa1NAc-T1およびNGa1NAc-T2の酵素活性における2価陽イオン要求性の検討結果

【0097】

【表5】

NGa1NAc-T1

2価陽イオン等	放射活性取込み量 (A)	ブランク (B)	(A) - (B)
MnCl ₂	519.47	263.21	256.26
MgCl ₂	256.36	263.21	-6.85
ZnCl ₂	210.29	263.21	-52.92
CaCl ₂	230.78	263.21	-32.43
CuCl ₂	278.77	263.21	15.56
CoCl ₂	240.91	263.21	-22.30
CdSO ₄	203.39	263.21	-59.82
EDTA	242.38	263.21	-20.83

(dpm)

NGa1NAc-T2

2価陽イオン等	放射活性取込み量 (A)	ブランク (B)	(A) - (B)
MnCl ₂	1484.43	263.21	1221.22
MgCl ₂	3124.16	263.21	2860.95
ZnCl ₂	187.59	263.21	-75.62
CaCl ₂	217.83	263.21	-45.38
CuCl ₂	218.35	263.21	-44.86
CoCl ₂	1130.63	263.21	867.42
CdSO ₄	217.92	263.21	-45.29
EDTA	235.28	263.21	-27.93

(dpm)

【0098】

実施例3

ヒト種々組織での発現解析

ヒト正常組織のcDNAを用いて、定量的PCRにより該遺伝子の発現量を定量した。正常組織のcDNAとしてはクロンテック社の総RNAから逆転写を行ったものを使用し、株化細胞に関しては総RNAを抽出し、常法によりcDNAを作製し使用した。NGa1NAc-T1の定量的発現解析に使用したプライマーはK12-F3 (5' -ctg gtg gat ttc gag agc ga-3' (配列番号20))、K12-R3 (5' -tgc cgt cca gga tgt tgg-3' (配列番号21))、プローブはK12-MGB3 (5' -gcg gta gag gac gcc-3' (配列番号22))

)である。また、NG a l N A c - T 2 の定量的発現解析に使用したプライマーはK 1 3 - F 3 (5' -atc gtc atc act gac tat agc agt ga-3' (配列番号 2 3))、K 1 3 - R 3 (5' -gaa tgg cat cga tga ctc cag-3' (配列番号 2 4))、プローブはK 1 3 - M G B 3 (5' -ctc gtg aag gac ccg ca-3' (配列番号 2 5))である。プローブにはアプライドバイオシステムズ社のマイナーグ ループバインダーを結合したものを使用した。酵素および反応液にはU n i v e r s a l P C R M a s t e r M i xを使用し、A B I P R I S M 7 7 0 0 S e q u e n c e D e t e c t i o n S y s t e m (ともにアプライドバイオシステムズ社)により反応液量 2 5 μ l で定量を行った。定量の標準遺伝子としてはグリセルアルデヒド-3-リン酸脱水素酵素 (G A P D H) 遺伝子を使用し、既知濃度の鋳型DNAにより定量の検量線を作成し該遺伝子の発現量の標準化を行った。また、NG a l N A c - T 1 およびNG a l N A c - T 2 の標準DNAとして、p F L A G - N G a l N A c - T 1 およびp F L A G - N G a l N A c - T 2 を用いた。反応温度は5 0 $^{\circ}$ C 2 分、9 5 $^{\circ}$ C 1 0 分の後、9 5 $^{\circ}$ C 1 5 秒、6 0 $^{\circ}$ C 1 分を5 0 サイクル行った。結果を図 1 に示した。NG a l N A c - T 1 は神経系で、NG a l N A c - T 2 は胃や精巣でそれぞれ発現量が高いことが明らかになった。

【0099】

実施例 4

ヒト癌組織での発現解析

ヒト肺癌組織と同患者の肺正常組織での両遺伝子の発現を解析した。方法は実施例 3 と同様である。ただし、コントロール遺伝子として β -アクチン遺伝子を用い、定量にはアプライドバイオシステムズ社のPre-Developed TaqMan Assay Reagents Endogenous Control Human Beta-actinを使用した。その結果、殆どの患者で発現が上昇していた (図 2) 。この結果から、両遺伝子は少なくとも肺癌のマーカーとして用いることが可能であることが明らかになった。

【 0 1 0 0 】

【配列表】

SEQUENCE LISTING

<110> NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY
AMERSHAM BIOSCIENCES KK

<120> Novel acetylgalactosamine transferases and nucleic acids encoding
the same

<130> A-987-26

<160> 25

<210> 1

<211> 1039

<212> PRT

<213> Homo sapiens

<400> 1

Met Pro Arg Leu Pro Val Lys Lys Ile Arg Lys Gln Met Lys Leu Leu

1 5 10 15

Leu Leu Leu Leu Leu Leu Ser Cys Ala Ala Trp Leu Thr Tyr Val His

20 25 30

Leu Gly Leu Val Arg Gln Gly Arg Ala Leu Arg Gln Arg Leu Gly Tyr

35 40 45

Gly Arg Asp Gly Glu Lys Leu Thr Ser Glu Thr Asp Gly Arg Gly Val

50

55

60

His Ala Ala Pro Ser Thr Gln Arg Ala Glu Asp Ser Ser Glu Ser Arg

65

70

75

80

Glu Glu Glu Gln Ala Pro Glu Gly Arg Asp Leu Asp Met Leu Phe Pro

85

90

95

Gly Gly Ala Gly Arg Leu Pro Leu Asn Phe Thr His Gln Thr Pro Pro

100

105

110

Trp Arg Glu Glu Tyr Lys Gly Gln Val Asn Leu His Val Phe Glu Asp

115

120

125

Trp Cys Gly Gly Ala Val Gly His Leu Arg Arg Asn Leu His Phe Pro

130

135

140

Leu Phe Pro His Thr Arg Thr Thr Val Lys Lys Leu Ala Val Ser Pro

145

150

155

160

Lys Trp Lys Asn Tyr Gly Leu Arg Ile Phe Gly Phe Ile His Pro Ala

165

170

175

Arg Asp Gly Asp Val Gln Phe Ser Val Ala Ser Asp Asp Asn Ser Glu

180

185

190

Phe Trp Leu Ser Leu Asp Glu Ser Pro Ala Ala Ala Gln Leu Val Ala

195

200

205

Phe Val Gly Lys Thr Gly Ser Glu Trp Thr Ala Pro Gly Glu Phe Thr
210 215 220

Lys Phe Ser Ser Gln Val Ser Lys Pro Arg Arg Leu Met Ala Ser Arg
225 230 235 240

Arg Tyr Tyr Phe Glu Leu Leu His Lys Gln Asp Asp Arg Gly Ser Asp
245 250 255

His Val Glu Val Gly Trp Arg Ala Phe Leu Pro Gly Leu Lys Phe Glu
260 265 270

Val Ile Ser Ser Ala His Ile Ser Leu Tyr Thr Asp Glu Ser Ala Leu
275 280 285

Lys Met Asp His Val Ala His Val Pro Gln Ser Pro Ala Ser His Val
290 295 300

Gly Gly Arg Pro Pro Gln Glu Glu Thr Ser Ala Asp Met Leu Arg Pro
305 310 315 320

Asp Pro Arg Asp Thr Phe Phe Leu Thr Pro Arg Met Glu Ser Ser Ser
325 330 335

Leu Glu Asn Val Leu Glu Pro Cys Ala Tyr Ala Pro Thr Tyr Val Val
340 345 350

Lys Asp Phe Pro Ile Ala Arg Tyr Gln Gly Leu Gln Phe Val Tyr Leu
355 360 365

Ser Phe Val Tyr Pro Asn Asp Tyr Thr Arg Leu Thr His Met Glu Thr
370 375 380

Asp Asn Lys Cys Phe Tyr Arg Glu Ser Pro Leu Tyr Leu Glu Arg Phe
385 390 395 400

Gly Phe Tyr Lys Tyr Met Lys Met Asp Lys Glu Glu Gly Asp Glu Asp
405 410 415

Glu Glu Asp Glu Val Gln Arg Arg Ala Phe Leu Phe Leu Asn Pro Asp
420 425 430

Asp Phe Leu Asp Asp Glu Asp Glu Gly Glu Leu Leu Asp Ser Leu Glu
435 440 445

Pro Thr Glu Ala Ala Pro Pro Arg Ser Gly Pro Gln Ser Pro Ala Pro
450 455 460

Ala Ala Pro Ala Gln Pro Gly Ala Thr Leu Ala Pro Pro Thr Pro Pro
465 470 475 480

Arg Pro Arg Asp Gly Gly Thr Pro Arg His Ser Arg Ala Leu Ser Trp
485 490 495

Ala Ala Arg Ala Ala Arg Pro Leu Pro Leu Phe Leu Gly Arg Ala Pro
500 505 510

Pro Pro Arg Pro Ala Val Glu Gln Pro Pro Pro Lys Val Tyr Val Thr

515

520

525

Arg Val Arg Pro Gly Gln Arg Ala Ser Pro Arg Ala Pro Ala Pro Arg

530

535

540

Ala Pro Trp Pro Pro Phe Pro Gly Val Phe Leu His Pro Arg Pro Leu

545

550

555

560

Pro Arg Val Gln Leu Arg Ala Pro Pro Arg Pro Pro Arg Pro His Gly

565

570

575

Arg Arg Thr Gly Gly Pro Gln Ala Thr Gln Pro Arg Pro Pro Ala Arg

580

585

590

Ala Gln Ala Thr Gln Gly Gly Arg Glu Gly Gln Ala Arg Thr Leu Gly

595

600

605

Pro Ala Ala Pro Thr Val Asp Ser Asn Leu Ser Ser Glu Ala Arg Pro

610

615

620

Val Thr Ser Phe Leu Ser Leu Ser Gln Val Ser Gly Pro Gln Leu Pro

625

630

635

640

Gly Glu Gly Glu Glu Glu Glu Glu Gly Glu Asp Asp Gly Ala Pro Gly

645

650

655

Asp Glu Ala Ala Ser Glu Asp Ser Glu Glu Ala Ala Gly Pro Ala Leu

660

665

670

Gly Arg Trp Arg Glu Asp Ala Ile Asp Trp Gln Arg Thr Phe Ser Val
675 680 685

Gly Ala Val Asp Phe Glu Leu Leu Arg Ser Asp Trp Asn Asp Leu Arg
690 695 700

Cys Asn Val Ser Gly Asn Leu Gln Leu Pro Glu Ala Glu Ala Val Asp
705 710 715 720

Val Thr Ala Gln Tyr Met Glu Arg Leu Asn Ala Arg His Gly Gly Arg
725 730 735

Phe Ala Leu Leu Arg Ile Val Asn Val Glu Lys Arg Arg Asp Ser Ala
740 745 750

Arg Gly Ser Arg Phe Leu Leu Glu Leu Glu Leu Gln Glu Arg Gly Gly
755 760 765

Gly Arg Leu Arg Leu Ser Glu Tyr Val Phe Leu Arg Leu Pro Gly Ala
770 775 780

Arg Val Gly Asp Ala Asp Gly Glu Ser Pro Glu Pro Ala Pro Ala Ala
785 790 795 800

Ser Val Arg Pro Asp Gly Arg Pro Glu Leu Cys Arg Pro Leu Arg Leu
805 810 815

Ala Trp Arg Gln Asp Val Met Val His Phe Ile Val Pro Val Lys Asn
820 825 830

Gln Ala Arg Trp Val Ala Gln Phe Leu Ala Asp Met Ala Ala Leu His
835 840 845

Ala Arg Thr Gly Asp Ser Arg Phe Ser Val Val Leu Val Asp Phe Glu
850 855 860

Ser Glu Asp Met Asp Val Glu Arg Ala Leu Arg Ala Ala Arg Leu Pro
865 870 875 880

Arg Tyr Gln Tyr Leu Arg Arg Thr Gly Asn Phe Glu Arg Ser Ala Gly
885 890 895

Leu Gln Ala Gly Val Asp Ala Val Glu Asp Ala Ser Ser Ile Val Phe
900 905 910

Leu Cys Asp Leu His Ile His Phe Pro Pro Asn Ile Leu Asp Gly Ile
915 920 925

Arg Lys His Cys Val Glu Gly Arg Leu Ala Phe Ala Pro Val Val Met
930 935 940

Arg Leu Ser Cys Gly Ser Ser Pro Arg Asp Pro His Gly Tyr Trp Glu
945 950 955 960

Val Asn Gly Phe Gly Leu Phe Gly Ile Tyr Lys Ser Asp Phe Asp Arg
965 970 975

Val Gly Gly Met Asn Thr Glu Glu Phe Arg Asp Gln Trp Gly Gly Glu

980

985

990

Asp Trp Glu Leu Leu Asp Arg Val Leu Gln Ala Gly Leu Glu Val Glu

995

1000

1005

Arg Leu Arg Leu Arg Asn Phe Tyr His His Tyr His Ser Lys Arg Gly

1010

1015

1020

Met Trp Ser Val Arg Ser Arg Lys Gly Ser Arg Thr Gly Ala Ser

1025

1030

1035

1039

<210> 2

<211> 3120

<212> DNA

<213> Homo sapiens

<400> 2

atg ccg cgg ctc ccg gtg aag aag atc cgt aag cag atg aag ctg ctg 48

Met Pro Arg Leu Pro Val Lys Lys Ile Arg Lys Gln Met Lys Leu Leu

1

5

10

15

ctg ctg ctg ctg ctg ctg agc tgc gcc gcg tgg ctc acc tac gtg cac 96

Leu Leu Leu Leu Leu Leu Ser Cys Ala Ala Trp Leu Thr Tyr Val His

20

25

30

ctg ggc ctg gtg cgc cag gga cgc gcg ctg cgc cag cgc ctg ggc tac 144

Leu Gly Leu Val Arg Gln Gly Arg Ala Leu Arg Gln Arg Leu Gly Tyr

35

40

45

ggg cga gat ggt gag aag ctg acc agt gag acc gac ggc cgg ggg gtc 192
Gly Arg Asp Gly Glu Lys Leu Thr Ser Glu Thr Asp Gly Arg Gly Val
50 55 60

cac gct gcg cca tcc aca cag agg gct gag gac tcc agt gag agc cgt 240
His Ala Ala Pro Ser Thr Gln Arg Ala Glu Asp Ser Ser Glu Ser Arg
65 70 75 80

gaa gag gag caa gcg ccc gaa ggt cgg gac cta gac atg ctg ttt cct 288
Glu Glu Glu Gln Ala Pro Glu Gly Arg Asp Leu Asp Met Leu Phe Pro
85 90 95

ggg ggg gct ggg agg ctg cca ctg aac ttc acc cat cag aca ccc cca 336
Gly Gly Ala Gly Arg Leu Pro Leu Asn Phe Thr His Gln Thr Pro Pro
100 105 110

tgg cgg gag gag tac aag ggg cag gtg aac ctg cac gtg ttt gag gac 384
Trp Arg Glu Glu Tyr Lys Gly Gln Val Asn Leu His Val Phe Glu Asp
115 120 125

tgg tgt ggg ggc gcc gtg ggc cac ctg agg agg aac ctg cac ttc ccg 432
Trp Cys Gly Gly Ala Val Gly His Leu Arg Arg Asn Leu His Phe Pro
130 135 140

ctg ttc cct cat acg cgc acc acc gtg aag aag ttg gcc gtg tcc ccc 480
Leu Phe Pro His Thr Arg Thr Thr Val Lys Lys Leu Ala Val Ser Pro
145 150 155 160

aag tgg aag aac tat gga ctc cgt att ttt ggt ttc atc cac ccg gcg 528
Lys Trp Lys Asn Tyr Gly Leu Arg Ile Phe Gly Phe Ile His Pro Ala
165 170 175

agg gac gga gac gtc cag ttt tct gtg gcc tca gac gac aac tcg gag 567
Arg Asp Gly Asp Val Gln Phe Ser Val Ala Ser Asp Asp Asn Ser Glu
180 185 190

ttc tgg ctg agt ctg gac gag agc cct gct gct gcc cag ctt gtg gcc 624
Phe Trp Leu Ser Leu Asp Glu Ser Pro Ala Ala Ala Gln Leu Val Ala
195 200 205

ttt gtg ggc aag act ggc tcc gag tgg aca gcg cct gga gaa ttc acc 672
Phe Val Gly Lys Thr Gly Ser Glu Trp Thr Ala Pro Gly Glu Phe Thr
210 215 220

aag ttc agc tcc cag gtg tcc aag ccc agg cgg ctc atg gcc tcc cgg 720
Lys Phe Ser Ser Gln Val Ser Lys Pro Arg Arg Leu Met Ala Ser Arg
225 230 235 240

agg tac tac ttt gag ttg ctg cac aag cag gac gac cgc ggc tcg gac 768
Arg Tyr Tyr Phe Glu Leu Leu His Lys Gln Asp Asp Arg Gly Ser Asp
245 250 255

cac gtg gaa gtg ggc tgg cga gct ttc ctg ccc ggc ctg aag ttc gag 816
His Val Glu Val Gly Trp Arg Ala Phe Leu Pro Gly Leu Lys Phe Glu
260 265 270

gtc atc agc tct gct cac atc tcc ctg tac aca gat gag tca gcc ttg 864

Val Ile Ser Ser Ala His Ile Ser Leu Tyr Thr Asp Glu Ser Ala Leu
275 280 285

aag atg gac cac gtg gcg cac gtc ccc cag tct cca gcc agc cac gtg 912
Lys Met Asp His Val Ala His Val Pro Gln Ser Pro Ala Ser His Val
290 295 300

ggg ggg cgt ccg ccg cag gag gag acc agc gca gac atg ctg cgg cca 960
Gly Gly Arg Pro Pro Gln Glu Glu Thr Ser Ala Asp Met Leu Arg Pro
305 310 315 320

gat ccc agg gat acc ttt ttc ctc act cca cgc atg gaa tct tcg agc 1008
Asp Pro Arg Asp Thr Phe Phe Leu Thr Pro Arg Met Glu Ser Ser Ser
325 330 335

ctg gag aac gtg ctg gag ccc tgc gcc tac gcc ccc acc tac gtg gtc 1056
Leu Glu Asn Val Leu Glu Pro Cys Ala Tyr Ala Pro Thr Tyr Val Val
340 345 350

aag gac ttc ccg atc gcc aga tac cag ggc ctg caa ttt gtg tac ctg 1104
Lys Asp Phe Pro Ile Ala Arg Tyr Gln Gly Leu Gln Phe Val Tyr Leu
355 360 365

tcc ttc gtt tat ccc aac gac tac act cgc ctc acc cac atg gag acg 1152
Ser Phe Val Tyr Pro Asn Asp Tyr Thr Arg Leu Thr His Met Glu Thr
370 375 380

gac aac aag tgc ttc tac cgc gag tct ccg ctg tat ctg gag agg ttt 1200
Asp Asn Lys Cys Phe Tyr Arg Glu Ser Pro Leu Tyr Leu Glu Arg Phe

385 390 395 400

ggg ttc tat aaa tac atg aag atg gac aag gag gag ggg gat gag gat 1248

Gly Phe Tyr Lys Tyr Met Lys Met Asp Lys Glu Glu Gly Asp Glu Asp
405 410 415

gaa gaa gac gag gtg cag cgc cga gcc ttc ctc ttc ctc aac ccg gac 1296

Glu Glu Asp Glu Val Gln Arg Arg Ala Phe Leu Phe Leu Asn Pro Asp
420 425 430

gac ttc ctg gac gac gag gac gag ggg gag ctg ctc gac agc ctg gag 1344

Asp Phe Leu Asp Asp Glu Asp Glu Gly Glu Leu Leu Asp Ser Leu Glu
435 440 445

ccc acc gag gcg gcc ccg ccc agg agc ggc ccc cag tcc ccc gcc cca 1392

Pro Thr Glu Ala Ala Pro Pro Arg Ser Gly Pro Gln Ser Pro Ala Pro
450 455 460

gca gcc ccc gcc cag ccc gga gcc acc ctc gcc ccg ccg acc cct ccc 1440

Ala Ala Pro Ala Gln Pro Gly Ala Thr Leu Ala Pro Pro Thr Pro Pro
465 470 475 480

cgc ccc cgg gac ggg ggg acc ccc agg cac tcc cgg gcc ctg agc tgg 1488

Arg Pro Arg Asp Gly Gly Thr Pro Arg His Ser Arg Ala Leu Ser Trp
485 490 495

gcc gcc agg gcc gcc cgc cct ttg ccg ctc ttc ttg ggc cga gct ccg 1536

Ala Ala Arg Ala Ala Arg Pro Leu Pro Leu Phe Leu Gly Arg Ala Pro
500 505 510

ccc ccg cgc cct gca gtg gag cag ccg ccc cca aag gtg tac gtg acc 1584

Pro Pro Arg Pro Ala Val Glu Gln Pro Pro Pro Lys Val Tyr Val Thr

515

520

525

agg gtg cgg ccg gga cag cgg gca tcc ccc cgg gcc cca gcg ccg cgt 1632

Arg Val Arg Pro Gly Gln Arg Ala Ser Pro Arg Ala Pro Ala Pro Arg

530

535

540

gcg ccc tgg ccg ccc ttc cct ggc gtc ttc ctg cac ccc agg cct ctg 1680

Ala Pro Trp Pro Pro Phe Pro Gly Val Phe Leu His Pro Arg Pro Leu

545

550

555

560

ccc aga gtg cag ctg cgg gcg ccc cca cgc cca ccc cgg ccc cac ggc 1728

Pro Arg Val Gln Leu Arg Ala Pro Pro Arg Pro Pro Arg Pro His Gly

565

570

575

cgc agg acc ggc ggc ccc cag gcc aca cag ccg agg ccc cca gcc cgg 1776

Arg Arg Thr Gly Gly Pro Gln Ala Thr Gln Pro Arg Pro Pro Ala Arg

580

585

590

gcg cag gcc acc caa ggg ggc cgg gag ggc cag gcg cgc acg ctg gga 1824

Ala Gln Ala Thr Gln Gly Gly Arg Glu Gly Gln Ala Arg Thr Leu Gly

595

600

605

cct gcg gcg ccc aca gtg gac tca aac ttg tcc tcc gaa gcg cgg ccc 1872

Pro Ala Ala Pro Thr Val Asp Ser Asn Leu Ser Ser Glu Ala Arg Pro

610

615

620

gtg acc tcc ttc ctg agc ttg tcc cag gtg tcc ggg ccg cag ctg ccc 1920
Val Thr Ser Phe Leu Ser Leu Ser Gln Val Ser Gly Pro Gln Leu Pro
625 630 635 640

ggg gag ggc gaa gag gag gag gaa ggg gag gac gat ggg gcc ccg ggc 1968
Gly Glu Gly Glu Glu Glu Glu Glu Gly Glu Asp Asp Gly Ala Pro Gly
645 650 655

gac gag gcc gcg tcg gag gac agc gag gag gcc gcg ggc ccg gcg ctc 2016
Asp Glu Ala Ala Ser Glu Asp Ser Glu Glu Ala Ala Gly Pro Ala Leu
660 665 670

gga cgc tgg cgt gag gac gcc atc gac tgg cag cgc acg ttc agc gtg 2064
Gly Arg Trp Arg Glu Asp Ala Ile Asp Trp Gln Arg Thr Phe Ser Val
675 680 685

ggc gcc gtg gac ttc gag ctg ctg cgc tcg gac tgg aac gac ctg cga 2112
Gly Ala Val Asp Phe Glu Leu Leu Arg Ser Asp Trp Asn Asp Leu Arg
690 695 700

tgc aac gtt tcg ggg aac ctg cag ctg ccg gag gcg gag gcc gtg gac 2160
Cys Asn Val Ser Gly Asn Leu Gln Leu Pro Glu Ala Glu Ala Val Asp
705 710 715 720

gtg acc gct cag tac atg gag cgg ctg aac gcg cgc cac ggc ggg cgc 2208
Val Thr Ala Gln Tyr Met Glu Arg Leu Asn Ala Arg His Gly Gly Arg
725 730 735

ttc gcg ctt ctg cgc atc gtg aac gtg gag aag cgc cgg gac tcg gcg 2256

Phe Ala Leu Leu Arg Ile Val Asn Val Glu Lys Arg Arg Asp Ser Ala
740 745 750

cga ggg agt cgc ttc ctg ctg gag ctg gag ctg cag gag cgc ggg ggc 2304
Arg Gly Ser Arg Phe Leu Leu Glu Leu Glu Leu Gln Glu Arg Gly Gly
755 760 765

ggc cgc ctg cga ctg tcc gag tac gtc ttc ctg cgg ctg ccg gga gcc 2352
Gly Arg Leu Arg Leu Ser Glu Tyr Val Phe Leu Arg Leu Pro Gly Ala
770 775 780

cgc gta ggg gat gca gac gga gaa agt ccc gaa ccc gct ccc gcc gcc 2400
Arg Val Gly Asp Ala Asp Gly Glu Ser Pro Glu Pro Ala Pro Ala Ala
785 790 795 800

tcc gtg cgc ccc gac ggc cgc ccc gag ctc tgc cgg cca ctg cgc ctg 2448
Ser Val Arg Pro Asp Gly Arg Pro Glu Leu Cys Arg Pro Leu Arg Leu
805 810 815

gcc tgg cgc cag gac gtg atg gtt cac ttc atc gtg cca gtg aaa aac 2496
Ala Trp Arg Gln Asp Val Met Val His Phe Ile Val Pro Val Lys Asn
820 825 830

cag gca cgg tgg gtg gca cag ttc ctg gcg gac atg gct gcg ctg cac 2544
Gln Ala Arg Trp Val Ala Gln Phe Leu Ala Asp Met Ala Ala Leu His
835 840 845

gcg cgc acc ggg gac tcg cgt ttc agc gtc gtc ctg gtg gat ttc gag 2592
Ala Arg Thr Gly Asp Ser Arg Phe Ser Val Val Leu Val Asp Phe Glu

850

855

860

agc gag gat atg gac gtg gag cgg gcc ctg cgc gcc gcg cgc ctg ccc 2640

Ser Glu Asp Met Asp Val Glu Arg Ala Leu Arg Ala Ala Arg Leu Pro

865

870

875

880

cgg tac cag tac ctg aga cga acc ggg aac ttc gag cgc tcc gcc ggg 2688

Arg Tyr Gln Tyr Leu Arg Arg Thr Gly Asn Phe Glu Arg Ser Ala Gly

885

890

895

ctg cag gcg gga gtg gac gcg gta gag gac gcc agc agc atc gtg ttc 2736

Leu Gln Ala Gly Val Asp Ala Val Glu Asp Ala Ser Ser Ile Val Phe

900

905

910

ctc tgc gac ctg cac atc cac ttc cca ccc aac atc ctg gac ggc atc 2784

Leu Cys Asp Leu His Ile His Phe Pro Pro Asn Ile Leu Asp Gly Ile

915

920

925

cgc aag cac tgc gtg gag ggc agg ctg gcc ttc gcg ccc gtg gtc atg 2832

Arg Lys His Cys Val Glu Gly Arg Leu Ala Phe Ala Pro Val Val Met

930

935

940

cgc ctg agc tgc ggg agc tcg ccc cgg gac ccc cac ggt tac tgg gag 2880

Arg Leu Ser Cys Gly Ser Ser Pro Arg Asp Pro His Gly Tyr Trp Glu

945

950

955

960

gtg aac ggc ttt ggc ctt ttt ggg atc tac aag tcg gac ttt gac cgg 2928

Val Asn Gly Phe Gly Leu Phe Gly Ile Tyr Lys Ser Asp Phe Asp Arg

965

970

975

gtt gga gga atg aac acg gag gag ttc cga gac cag tgg ggg ggt gaa 2976

Val Gly Gly Met Asn Thr Glu Glu Phe Arg Asp Gln Trp Gly Gly Glu

980

985

990

gac tgg gag ctc ctg gac agg gtc ctg cag gca ggg ctg gag gtg gag 3024

Asp Trp Glu Leu Leu Asp Arg Val Leu Gln Ala Gly Leu Glu Val Glu

995

1000

1005

cgg ctc cga ctg cgg aat ttc tat cac cac tac cac tcc aag agg ggc 3072

Arg Leu Arg Leu Arg Asn Phe Tyr His His Tyr His Ser Lys Arg Gly

1010

1015

1020

atg tgg agc gtc cgc agc agg aag ggc tct cgc acg ggg gcg tct tga 3120

Met Trp Ser Val Arg Ser Arg Lys Gly Ser Arg Thr Gly Ala Ser

1025

1030

1035

1039

<210> 3

<211> 998

<212> PRT

<213> Homo sapiens

<400> 3

Met Gly Ser Pro Arg Ala Ala Arg Pro Pro Leu Leu Leu Arg Pro Val

1

5

10

15

Lys Leu Leu Arg Arg Arg Phe Arg Leu Leu Leu Ala Leu Ala Val Val

20

25

30

Ser Val Gly Leu Trp Thr Leu Tyr Leu Glu Leu Val Ala Ser Ala Gln
35 40 45

Val Gly Gly Asn Pro Leu Asn Arg Arg Tyr Gly Ser Trp Arg Glu Leu
50 55 60

Ala Lys Ala Leu Ala Ser Arg Asn Ile Pro Ala Val Asp Pro His Leu
65 70 75 80

Gln Phe Tyr His Pro Gln Arg Leu Ser Leu Glu Asp His Asp Ile Asp
85 90 95

Gln Gly Val Ser Ser Asn Ser Ser Tyr Leu Lys Trp Asn Lys Pro Val
100 105 110

Pro Trp Leu Ser Glu Phe Arg Gly Arg Ala Asn Leu His Val Phe Glu
115 120 125

Asp Trp Cys Gly Ser Ser Ile Gln Gln Leu Arg Arg Asn Leu His Phe
130 135 140

Pro Leu Tyr Pro His Ile Arg Thr Thr Leu Arg Lys Leu Ala Val Ser
145 150 155 160

Pro Lys Trp Thr Asn Tyr Gly Leu Arg Ile Phe Gly Tyr Leu His Pro
165 170 175

Phe Thr Asp Gly Lys Ile Gln Phe Ala Ile Ala Ala Asp Asp Asn Ala

180

185

190

Glu Phe Trp Leu Ser Leu Asp Asp Gln Val Ser Gly Leu Gln Leu Leu

195

200

205

Ala Ser Val Gly Lys Thr Gly Lys Glu Trp Thr Ala Pro Gly Glu Phe

210

215

220

Gly Lys Phe Arg Ser Gln Ile Ser Lys Pro Val Ser Leu Ser Ala Ser

225

230

235

240

His Arg Tyr Tyr Phe Glu Val Leu His Lys Gln Asn Glu Glu Gly Thr

245

250

255

Asp His Val Glu Val Ala Trp Arg Arg Asn Asp Pro Gly Ala Lys Phe

260

265

270

Thr Ile Ile Asp Ser Leu Ser Leu Ser Leu Phe Thr Asn Glu Thr Phe

275

280

285

Leu Gln Met Asp Glu Val Gly His Ile Pro Gln Thr Ala Ala Ser His

290

295

300

Val Asp Ser Ser Asn Ala Leu Pro Arg Asp Glu Gln Pro Pro Ala Asp

305

310

315

320

Met Leu Arg Pro Asp Pro Arg Asp Thr Leu Tyr Arg Val Pro Leu Ile

325

330

335

Pro Lys Ser His Leu Arg His Val Leu Pro Asp Cys Pro Tyr Lys Pro
340 345 350

Ser Tyr Leu Val Asp Gly Leu Pro Leu Gln Arg Tyr Gln Gly Leu Arg
355 360 365

Phe Val His Leu Ser Phe Val Tyr Pro Asn Asp Tyr Thr Arg Leu Ser
370 375 380

His Met Glu Thr His Asn Lys Cys Phe Tyr Gln Glu Asn Ala Tyr Tyr
385 390 395 400

Gln Asp Arg Phe Ser Phe Gln Glu Tyr Ile Arg Ile Asp Gln Pro Glu
405 410 415

Lys Gln Gly Leu Glu Gln Pro Gly Phe Glu Glu Asn Leu Leu Glu Glu
420 425 430

Ser Gln Tyr Gly Glu Val Ala Glu Glu Thr Pro Ala Ser Asn Asn Gln
435 440 445

Asn Ala Arg Met Leu Glu Gly Arg Gln Thr Pro Ala Ser Thr Leu Glu
450 455 460

Gln Asp Ala Thr Asp Tyr Arg Leu Arg Ser Leu Arg Lys Leu Leu Ala
465 470 475 480

Gln Pro Arg Glu Gly Leu Leu Ala Pro Phe Ser Lys Arg Asn Ser Thr
485 490 495

Ala Ser Phe Pro Gly Arg Thr Ser His Ile Pro Val Gln Gln Pro Glu
500 505 510

Lys Arg Lys Gln Lys Pro Ser Pro Glu Pro Ser Gln Asp Ser Pro His
515 520 525

Ser Asp Lys Trp Pro Pro Gly His Pro Val Lys Asn Leu Pro Gln Met
530 535 540

Arg Gly Pro Arg Pro Arg Pro Ala Gly Asp Ser Pro Arg Lys Thr Gln
545 550 555 560

Trp Leu Asn Gln Val Glu Ser Tyr Ile Ala Glu Gln Arg Arg Gly Asp
565 570 575

Arg Met Arg Pro Gln Ala Pro Gly Arg Gly Trp His Gly Glu Glu Glu
580 585 590

Val Val Ala Ala Ala Gly Gln Glu Gly Gln Val Glu Gly Glu Glu Glu
595 600 605

Gly Glu Glu Glu Glu Glu Glu Glu Asp Met Ser Glu Val Phe Glu Tyr
610 615 620

Val Pro Val Phe Asp Pro Val Val Asn Trp Asp Gln Thr Phe Ser Ala
625 630 635 640

Arg Asn Leu Asp Phe Gln Ala Leu Arg Thr Asp Trp Ile Asp Leu Ser

645

650

655

Cys Asn Thr Ser Gly Asn Leu Leu Leu Pro Glu Gln Glu Ala Leu Glu

660

665

670

Val Thr Arg Val Phe Leu Lys Lys Leu Asn Gln Arg Ser Arg Gly Arg

675

680

685

Tyr Gln Leu Gln Arg Ile Val Asn Val Glu Lys Arg Gln Asp Gln Leu

690

695

700

Arg Gly Gly Arg Tyr Leu Leu Glu Leu Glu Leu Leu Glu Gln Gly Gln

705

710

715

720

Arg Val Val Arg Leu Ser Glu Tyr Val Ser Ala Arg Gly Trp Gln Gly

725

730

735

Ile Asp Pro Ala Gly Gly Glu Glu Val Glu Ala Arg Asn Leu Gln Gly

740

745

750

Leu Val Trp Asp Pro His Asn Arg Arg Arg Gln Val Leu Asn Thr Arg

755

760

765

Ala Gln Glu Pro Lys Leu Cys Trp Pro Gln Gly Phe Ser Trp Ser His

770

775

780

Arg Ala Val Val His Phe Val Val Pro Val Lys Asn Gln Ala Arg Trp

785

790

795

800

Val Gln Gln Phe Ile Lys Asp Met Glu Asn Leu Phe Gln Val Thr Gly
805 810 815

Asp Pro His Phe Asn Ile Val Ile Thr Asp Tyr Ser Ser Glu Asp Met
820 825 830

Asp Val Glu Met Ala Leu Lys Arg Ser Lys Leu Arg Ser Tyr Gln Tyr
835 840 845

Val Lys Leu Ser Gly Asn Phe Glu Arg Ser Ala Gly Leu Gln Ala Gly
850 855 860

Ile Asp Leu Val Lys Asp Pro His Ser Ile Ile Phe Leu Cys Asp Leu
865 870 875 880

His Ile His Phe Pro Ala Gly Val Ile Asp Ala Ile Arg Lys His Cys
885 890 895

Val Glu Gly Lys Met Ala Phe Ala Pro Met Val Met Arg Leu His Cys
900 905 910

Gly Ala Thr Pro Gln Trp Pro Glu Gly Tyr Trp Glu Val Asn Gly Phe
915 920 925

Gly Leu Leu Gly Ile Tyr Lys Ser Asp Leu Asp Arg Ile Gly Gly Met
930 935 940

Asn Thr Lys Glu Phe Arg Asp Arg Trp Gly Gly Glu Asp Trp Glu Leu

945

950

955

960

Leu Asp Arg Ile Leu Gln Ala Gly Leu Asp Val Glu Arg Leu Ser Leu

965

970

975

Arg Asn Phe Phe His His Phe His Ser Lys Arg Gly Met Trp Ser Arg

980

985

990

Arg Gln Met Lys Thr Leu

995

998

<210> 4

<211> 2997

<212> DNA

<213> Homo sapiens

<400> 4

atg ggg agc ccc cgg gcc gcg cgg ccc ccg ctg ctc ctg cgc ccg gtg 48

Met Gly Ser Pro Arg Ala Ala Arg Pro Pro Leu Leu Leu Arg Pro Val

1

5

10

15

aag ctg ctg cgg agg cgc ttc cgg ctg ctg ctg gcg ctc gcc gtg gtg 96

Lys Leu Leu Arg Arg Arg Phe Arg Leu Leu Leu Ala Leu Ala Val Val

20

25

30

tct gtg ggg ctc tgg act ctg tat ctg gaa ctg gtg gcg tcg gcc cag 144

Ser Val Gly Leu Trp Thr Leu Tyr Leu Glu Leu Val Ala Ser Ala Gln

35

40

45

gtc ggc ggg aac ccc ctg aac cgg agg tac ggc agc tgg aga gaa cta	192
Val Gly Gly Asn Pro Leu Asn Arg Arg Tyr Gly Ser Trp Arg Glu Leu	
50 55 60	
gcc aag gct ctg gcc agc agg aac att cca gct gtg gat cca cac ctc	240
Ala Lys Ala Leu Ala Ser Arg Asn Ile Pro Ala Val Asp Pro His Leu	
65 70 75 80	
cag ttc tac cat ccc cag agg ctg agc ctc gag gac cac gac att gac	288
Gln Phe Tyr His Pro Gln Arg Leu Ser Leu Glu Asp His Asp Ile Asp	
85 90 95	
caa ggg gtg agc agt aac agc agc tac ttg aag tgg aac aag cct gtc	336
Gln Gly Val Ser Ser Asn Ser Ser Tyr Leu Lys Trp Asn Lys Pro Val	
100 105 110	
ccc tgg ctc tca gag ttc cgg ggc cgt gcc aac ctg cat gtg ttt gaa	384
Pro Trp Leu Ser Glu Phe Arg Gly Arg Ala Asn Leu His Val Phe Glu	
115 120 125	
gac tgg tgt ggc agc tct atc cag cag ctc agg agg aac ctg cat ttc	432
Asp Trp Cys Gly Ser Ser Ile Gln Gln Leu Arg Arg Asn Leu His Phe	
130 135 140	
cca ctg tac ccc cat att cgc aca acc ctg agg aag ctt gct gtg tcc	480
Pro Leu Tyr Pro His Ile Arg Thr Thr Leu Arg Lys Leu Ala Val Ser	
145 150 155 160	

ccc aaa tgg acc aac tat ggc ctc cgc atc ttt ggc tac ctg cac ccc 528
 Pro Lys Trp Thr Asn Tyr Gly Leu Arg Ile Phe Gly Tyr Leu His Pro
 165 170 175

ttt act gat ggg aaa atc cag ttt gcc att gct gca gat gac aac gcg 576
 Phe Thr Asp Gly Lys Ile Gln Phe Ala Ile Ala Ala Asp Asp Asn Ala
 180 185 190

gag ttc tgg ctg agc ctc gat gac cag gtc tca ggc ctc cag ctg ctg 624
 Glu Phe Trp Leu Ser Leu Asp Asp Gln Val Ser Gly Leu Gln Leu Leu
 195 200 205

gcc agt gtg ggc aag act gga aag gag tgg acc gcc ccg gga gag ttt 672
 Ala Ser Val Gly Lys Thr Gly Lys Glu Trp Thr Ala Pro Gly Glu Phe
 210 215 220

ggg aaa ttt cgg agc caa att tcc aag ccg gtg agc ctg tca gcc tcc 720
 Gly Lys Phe Arg Ser Gln Ile Ser Lys Pro Val Ser Leu Ser Ala Ser
 225 230 235 240

cac agg tac tac ttc gag gtg ctg cac aag cag aat gag gag ggc acc 768
 His Arg Tyr Tyr Phe Glu Val Leu His Lys Gln Asn Glu Glu Gly Thr
 245 250 255

gac cac gtg gaa gtt gca tgg cga cgg aac gac cct gga gcc aag ttc 816
 Asp His Val Glu Val Ala Trp Arg Arg Asn Asp Pro Gly Ala Lys Phe
 260 265 270

acc atc att gac tcc ctc tcc ctg tcc ctc ttc aca aat gag acg ttc 864

Thr Ile Ile Asp Ser Leu Ser Leu Ser Leu Phe Thr Asn Glu Thr Phe
275 280 285

cta cag atg gat gag gtg ggc cac atc cca cag aca gca gcc agc cac 912
Leu Gln Met Asp Glu Val Gly His Ile Pro Gln Thr Ala Ala Ser His
290 295 300

gtg gac tcc tcc aac gct ctt ccc agg gat gag cag ccg ccc gct gac 960
Val Asp Ser Ser Asn Ala Leu Pro Arg Asp Glu Gln Pro Pro Ala Asp
305 310 315 320

atg ctt cgg cct gac ccc cgg gac acc ctc tat cga gtg cct ctg atc 1008
Met Leu Arg Pro Asp Pro Arg Asp Thr Leu Tyr Arg Val Pro Leu Ile
325 330 335

ccc aag tcg cat ctc cgc cac gtc ctg cct gac tgt ccc tac aaa ccc 1056
Pro Lys Ser His Leu Arg His Val Leu Pro Asp Cys Pro Tyr Lys Pro
340 345 350

agc tat ctg gtg gat ggg ctt cct ctg cag cgc tac cag gga .ctc cgg 1104
Ser Tyr Leu Val Asp Gly Leu Pro Leu Gln Arg Tyr Gln Gly Leu Arg
355 360 365

ttt gtt cat ctg tct ttt gtt tac ccc aat gac tat acc cgc ctg agc 1152
Phe Val His Leu Ser Phe Val Tyr Pro Asn Asp Tyr Thr Arg Leu Ser
370 375 380

cac atg gag acc cac aat aaa tgt ttc tac cag gaa aac gcc tac tac 1200
His Met Glu Thr His Asn Lys Cys Phe Tyr Gln Glu Asn Ala Tyr Tyr

385 390 395 400

caa gac cgg ttc agc ttt cag gag tac atc agg att gac cag cct gag 1248

Gln Asp Arg Phe Ser Phe Gln Glu Tyr Ile Arg Ile Asp Gln Pro Glu

405 410 415

aag cag ggg ctg gag cag cca ggt ttt gag gaa aac ctt cta gaa gag 1296

Lys Gln Gly Leu Glu Gln Pro Gly Phe Glu Glu Asn Leu Leu Glu Glu

420 425 430

tcc cag tat ggg gaa gtg gca gag gag acc cct gcc tcc aac aac cag 1344

Ser Gln Tyr Gly Glu Val Ala Glu Glu Thr Pro Ala Ser Asn Asn Gln

435 440 445

aat gcc agg atg ctt gag gga aga cag aca cct gcc tcc acc ctg gag 1392

Asn Ala Arg Met Leu Glu Gly Arg Gln Thr Pro Ala Ser Thr Leu Glu

450 455 460

caa gat gcc act gac tac cgc ctc cga agc ctg cgg aaa ctc ctg gct 1440

Gln Asp Ala Thr Asp Tyr Arg Leu Arg Ser Leu Arg Lys Leu Leu Ala

465 470 475 480

cag ccc cgg gag ggc ctg ctg gcc ccc ttc tcc aag cgg aac tcc aca 1488

Gln Pro Arg Glu Gly Leu Leu Ala Pro Phe Ser Lys Arg Asn Ser Thr

485 490 495

gcg tcc ttc cca ggg agg acc agc cac att cca gtg cag cag cca gag 1536

Ala Ser Phe Pro Gly Arg Thr Ser His Ile Pro Val Gln Gln Pro Glu

500 505 510

aag agg aag caa aaa ccc agc cct gag ccc agc caa gat tca cct cat 1584

Lys Arg Lys Gln Lys Pro Ser Pro Glu Pro Ser Gln Asp Ser Pro His

515

520

525

tcc gac aag tgg cct cct ggg cac cct gtg aag aac ctg cct cag atg 1632

Ser Asp Lys Trp Pro Pro Gly His Pro Val Lys Asn Leu Pro Gln Met

530

535

540

agg ggg ccc agg ccc agg ccc gct ggt gac agc ccc agg aag act cag 1680

Arg Gly Pro Arg Pro Arg Pro Ala Gly Asp Ser Pro Arg Lys Thr Gln

545

550

555

560

tgg ctg aac cag gtg gag tcg tac atc gca gag cag aga cgg ggt gac 1728

Trp Leu Asn Gln Val Glu Ser Tyr Ile Ala Glu Gln Arg Arg Gly Asp

565

570

575

agg atg cgg cct cag gcc ccc gga agg ggc tgg cat ggg gag gag gaa 1776

Arg Met Arg Pro Gln Ala Pro Gly Arg Gly Trp His Gly Glu Glu Glu

580

585

590

gtg gtg gcg gcc gca ggc cag gaa gga caa gtg gag gga gag gaa gag 1824

Val Val Ala Ala Ala Gly Gln Glu Gly Gln Val Glu Gly Glu Glu Glu

595

600

605

ggg gaa gaa gag gag gag gaa gag gat atg agt gag gtg ttc gag tac 1872

Gly Glu Glu Glu Glu Glu Glu Glu Asp Met Ser Glu Val Phe Glu Tyr

610

615

620

gta cct gtg ttt gac ccg gta gta aac tgg gac cag acc ttc agt gcc 1920
Val Pro Val Phe Asp Pro Val Val Asn Trp Asp Gln Thr Phe Ser Ala
625 630 635 640

cgg aat ctc gac ttc caa gcc ctg agg act gac tgg atc gat ctg agc 1968.
Arg Asn Leu Asp Phe Gln Ala Leu Arg Thr Asp Trp Ile Asp Leu Ser
645 650 655

tgt aac aca tct ggc aac ctg ctg ctt cca gag cag gaa gct ctg gag 2016
Cys Asn Thr Ser Gly Asn Leu Leu Leu Pro Glu Gln Glu Ala Leu Glu
660 665 670

gtc acg cga gtc ttc ttg aag aag ctc aac cag agg agc cgg ggg agg 2064
Val Thr Arg Val Phe Leu Lys Lys Leu Asn Gln Arg Ser Arg Gly Arg
675 680 685

tac cag cta cag cgc att gtg aac gtg gaa aag cgt cag gac cag cta 2112
Tyr Gln Leu Gln Arg Ile Val Asn Val Glu Lys Arg Gln Asp Gln Leu
690 695 700

cgt ggg ggt cgc tac ctc ctg gag ctt gaa ctg ttg gaa caa ggc cag 2160
Arg Gly Gly Arg Tyr Leu Leu Glu Leu Glu Leu Leu Glu Gln Gly Gln
705 710 715 720

cgc gtg gtg cgg ctc tcg gag tat gtg tct gca cga ggc tgg cag ggc 2208
Arg Val Val Arg Leu Ser Glu Tyr Val Ser Ala Arg Gly Trp Gln Gly
725 730 735

atc gat cca gct ggt ggg gag gag gtc gag gcc cgg aac ctg caa ggc 2256

Ile Asp Pro Ala Gly Gly Glu Glu Val Glu Ala Arg Asn Leu Gln Gly
740 745 750

ctg gtc tgg gac cca cac aac cgt agg aga cag gtc ctg aat acc cgg 2304
Leu Val Trp Asp Pro His Asn Arg Arg Arg Gln Val Leu Asn Thr Arg
755 760 765

gcc caa gag ccc aag ctg tgc tgg cct cag ggt ttc tcc tgg agt cac 2352
Ala Gln Glu Pro Lys Leu Cys Trp Pro Gln Gly Phe Ser Trp Ser His
770 775 780

cga gcc gtg gtc cac ttc gtc gtg cct gtg aag aac cag gca cgc tgg 2400
Arg Ala Val Val His Phe Val Val Pro Val Lys Asn Gln Ala Arg Trp
785 790 795 800

gta cag caa ttc atc aaa gac atg gaa aac ctg ttc cag gtc acc ggt 2448
Val Gln Gln Phe Ile Lys Asp Met Glu Asn Leu Phe Gln Val Thr Gly
805 810 815

gac cca cac ttc aac atc gtc atc act gac tat agc agt gag gac atg 2496
Asp Pro His Phe Asn Ile Val Ile Thr Asp Tyr Ser Ser Glu Asp Met
820 825 830

gat gtt gag atg gca ctg aag agg tcc aag ctg cgg agc tac cag tac 2544
Asp Val Glu Met Ala Leu Lys Arg Ser Lys Leu Arg Ser Tyr Gln Tyr
835 840 845

gtg aag cta agt gga aac ttt gaa cgc tca gct gga ctt cag gct ggc 2592
Val Lys Leu Ser Gly Asn Phe Glu Arg Ser Ala Gly Leu Gln Ala Gly

850

855

860

ata gac ctc gtg aag gac ccg cac agc atc atc ttc ctc tgt gac ctc 2640
Ile Asp Leu Val Lys Asp Pro His Ser Ile Ile Phe Leu Cys Asp Leu
865 870 875 880

cac atc cac ttc cca gct gga gtc atc gat gcc att cgg aag cac tgt 2688
His Ile His Phe Pro Ala Gly Val Ile Asp Ala Ile Arg Lys His Cys
885 890 895

gtg gag gga aag atg gcc ttt gcc ccc atg gtg atg agg ctg cat tgt 2736
Val Glu Gly Lys Met Ala Phe Ala Pro Met Val Met Arg Leu His Cys
900 905 910

ggg gcc acc ccc cag tgg cct gag ggc tac tgg gag gtg aat ggg ttc 2784
Gly Ala Thr Pro Gln Trp Pro Glu Gly Tyr Trp Glu Val Asn Gly Phe
915 920 925

ggg ctg ctt ggc atc tac aag tct gac ctg gac agg att ggg ggc atg 2832
Gly Leu Leu Gly Ile Tyr Lys Ser Asp Leu Asp Arg Ile Gly Gly Met
930 935 940

aac acc aag gag ttc cga gac cgc tgg ggc ggg gaa gac tgg gag ctg 2880
Asn Thr Lys Glu Phe Arg Asp Arg Trp Gly Gly Glu Asp Trp Glu Leu
945 950 955 960

ctg gac agg ata ctc caa gcg ggc ctg gac gtg gag cgt ctc tcc ctc 2928
Leu Asp Arg Ile Leu Gln Ala Gly Leu Asp Val Glu Arg Leu Ser Leu
965 970 975

agg aat ttc ttc cat cat ttc cat tcc aag cga ggc atg tgg agc cgt 2976

Arg Asn Phe Phe His His Phe His Ser Lys Arg Gly Met Trp Ser Arg

980

985

990

cgc cag atg aag acg ctg tag

2997

Arg Gln Met Lys Thr Leu

995

998

<210> 5

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T1 cDNA

<400> 5

gctcctgcag ctccagctcc a 21

<210> 6

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T1 cDNA

<400> 6

aagcgactcc ctcgcgccga gt 22

<210> 7

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T1 cDNA

<400> 7

atgccgcggc tcccggtgaa gaag 24

<210> 8

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 8

ccacagttca agctccagga ggta 24

<210> 9

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 9

ctgacgcttt tccacgttca caat 24

<210> 10

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 10

caccccgctct ctgctctgcg at 22

<210> 11

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 11

gtcttcctgg ggctgtcacc a 21

<210> 12

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 12

cacatcatcc atctgtagga acgt 24

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 13

ctgtcgccct gcaacttcca cgt 23

<210> 14

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 14

aatgtcgtgg tcctcgaggc tca 23

<210> 15

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 15

gatggtagaa ctggaggtgt ggat 24

<210> 16

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T1 cDNA

<400> 16

cccaagcttc ggggggtcca cgctgcgcca t 31

<210> 17

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T1 cDNA

<400> 17

gctctagact caagacgccc ccgtgcgaga 30

<210> 18

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 18

ggaattcgag gtacggcagc tggagagaa 29

<210> 19

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for cloning GalNAc-T2 cDNA

<400> 19

acgcgtcgac ctacagcgtc ttcattctggc ga

<210> 20

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for amplifying GalNAc-T1 cDNA

<400> 20

ctggtggatt tcgagagcga 20

<210> 21

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for amplifying GalNAc-T1 cDNA

<400> 21

tgccgtccag gatgttgg 18

<210> 22

<211> 15

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide probe used in PCR for detecting GalNAc-T1 cDNA

<400> 22

gcggtagagg acgcc 15

<210> 23

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for amplifying GalNAc-T2 cDNA

<400> 23

atcgatcatca ctgactatag cagtga 26

<210> 24

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide primer used in PCR for amplifying GalNAc-T2 cDNA

<400> 24

gaatggcatc gatgactcca g 21

<210> 25

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide probe used in PCR for detecting GalNAc-T2 cDNA

<400> 25

ctcgtgaagg acccgca 17

参考文献

1. Sugita, M., S. Itonori, F. Inagaki and T. Hori, Characterization of two glucuronic acid-containing glycosphingolipids in larvae of the green-bottle fly, *Lucilia caesar*. *J. Biol. Chem.*, 1989. 264, p.15028-33
2. Helling, F., R.D. Dennis, B. Weske, G. Nores, J. Peter-Katalinic, U. Dabrowsli, H. Egge and H. Wiegandt, Glycosphingolipids in insects. The amphoteric moiety, N-acetylglucosamine-linked phosphoethanolamine, distinguishes a group of ceramide oligosaccharides from the pupae of *Calliphora vicina* (Insecta: Diptera). *Eur. J. Biochem.*, 1991. 200, p.409-21
3. Weisshaar, G., J. Hiyama, A.G. Renwick and M. Nimtz, NMR investigations of the N-linked oligosaccharides at individual glycosylation sites of human lutropin. *Eur. J. Biochem.*, 1991. 195, p.257-68
4. White, T., E.P. Bennet, K. Takio, T. Sorensen, N. Bonding and H. Clausen, Purification and cDNA cloning of a human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase. *J. Biol. Chem.*, 1995. 270, p.24156-65
5. Bennett, E.P., H. Hassan and H. Clausen, cDNA cloning and expression of a novel human UDP-N-acetyl-alpha-D-galactosamine. Polypeptide N-acetylgalactosaminyltransferase, GalNAc-t3. *J. Biol. Chem.*, 1996. 271, p.17006-12
6. Bennett, E.P., H. Hassan, U. Mandel, E. Mirgorodskaya, P. Roepstorff, J. Burchell, J. Taylor-Papadimitriou, M.A. Hollingsworth, G. Merks, A.G. van Kessel, H. Eiberg, R. Steffensen and H. Clausen, Cloning of a human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. *J. Biol. Chem.*, 1998. 273, p.30472-81

7. Bennett, E.P., H. Hassan, J. Mandel, M.A. Hollingsworth, N. Akisawa, Y. Ikematsu, G. Merckx, A.G. van Kessel, S. Olofsson and H. Clausen, Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-T3, designed GalNAc-T6. Evidence for genetic but not functional redundancy. *J. Biol. Chem.*, 1999. 274, p.25362-70
8. Bennett, E.P., H. Hassan, M.A. Hollingsworth and H. Clausen, A novel human UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. *FEBS Lett.*, 1999. 460, p.226-30
9. White, K. E., B. Lorenz, T. Meitinger, T.M. Strom and M.J. Econs, *Gene*, 2000. 246, p.347-56
10. Toba, S., M. Tenno, M. Konishi, T. Mikami, N. Itoh and A. Kurosaka, Brain-specific expression of a novel human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (GalNAc-T9). *Biochim. Biophys. Acta.*, 2000. 7, p.264-8
11. Nagata, Y., S. Yamashiro, J. Yodoi, K.O. Lloyd, H. Shiku and K. Furukawa, Expression cloning of beta 1,4 N-acetylgalactosaminyltransferase cDNAs that determine the expression of GM2 and GD2 gangliosides. *J. Biol. Chem.*, 1992. 269, p.12082-9
12. Yamamoto, F., J. Marken, T. Tsuji, T. White, H. Clausen and S. Hakomori, Cloning and characterization of DNA complementary to human UDP-GalNAc:Fuc alpha 1----2Gal alpha 1----3GalNAc transferase (histo-blood group A transferase) mRNA. *J. Biol. Chem.*, 1990. 265, p.1146-51
13. Xu, H., T. Storch, M. Yu, S.P. Elliott and D.B. Haslam, Characterization of the human Forssman synthetase gene. An evolving association between glycolipid synthesis and host-microbial interactions. *J. Biol. Chem.*, 1999. 274, p.29390-8
14. Guo, J.M., Y. Zhang, L. Cheng, H. Iwasaki, H. Wang, T. Kubota, K. Ta

chibana and H. Narimatsu, Molecular cloning and characterization of a novel member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family, pp-GalNAc-T12(1). FEBS Lett., 2002. 524, p.211-8

【0101】

【発明の効果】

本発明により、N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する酵素が単離され、その遺伝子の構造が明らかにされた。したがって、該酵素等の遺伝子工学的な生産や該酵素による糖鎖の生産、該遺伝子らに基づく疾患の診断が可能になった。

【図面の簡単な説明】

【図 1】

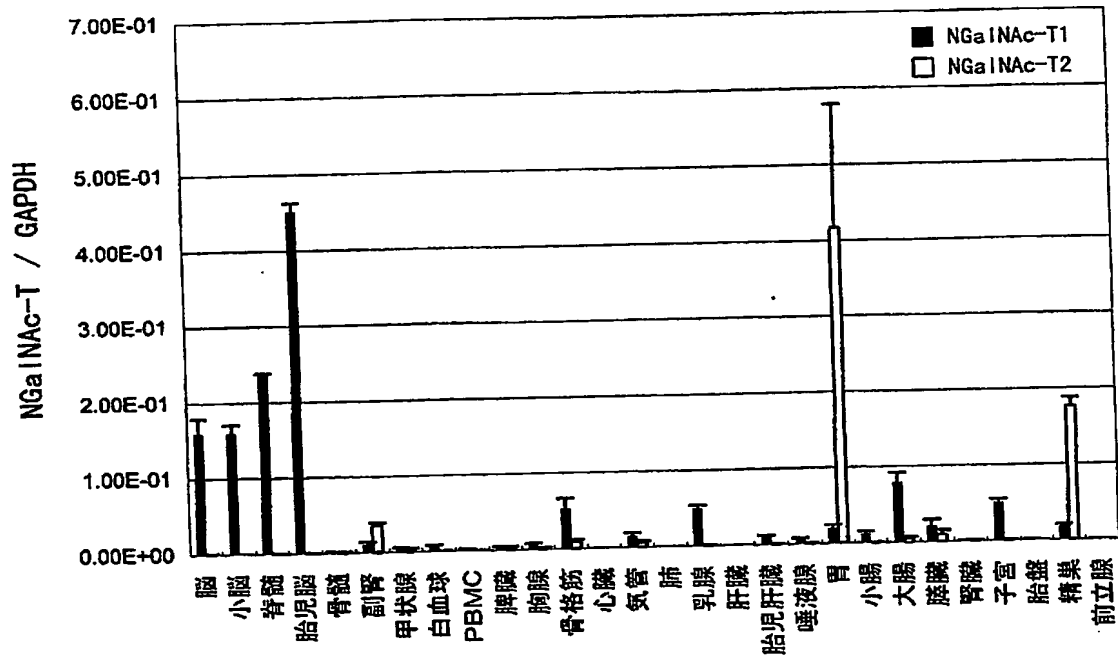
図1は、ヒト各種組織におけるNGa1NAc-T1またはNGa1NAc-T2遺伝子の発現量をリアルタイムPCR法により定量した結果を示す。縦軸はコントロールのグリセルアルデヒド-3-リン酸脱水素酵素 (GAPDH) 遺伝子の発現量に対するNGa1NAc-T1またはNGa1NAc-T2遺伝子の発現量の相対比を表す。各組織におけるNGa1NAc-T1遺伝子の発現は黒バーで示し、NGa1NAc-T2遺伝子の発現は白バーで示す。

【図 2】

図2は、ヒト肺癌組織と正常組織におけるNGa1NAc-T1 (パネルA) またはNGa1NAc-T2 (パネルB) 遺伝子の発現量をリアルタイムPCR法により定量した結果を示す。縦軸はコントロールのヒトβ-アクチン遺伝子の発現量に対するNGa1NAc-T1またはNGa1NAc-T2遺伝子の発現量の相対比を表す。横軸は、各患者の番号を示す。正常組織を白バー、癌組織を黒バーで示す。

【書類名】 図面

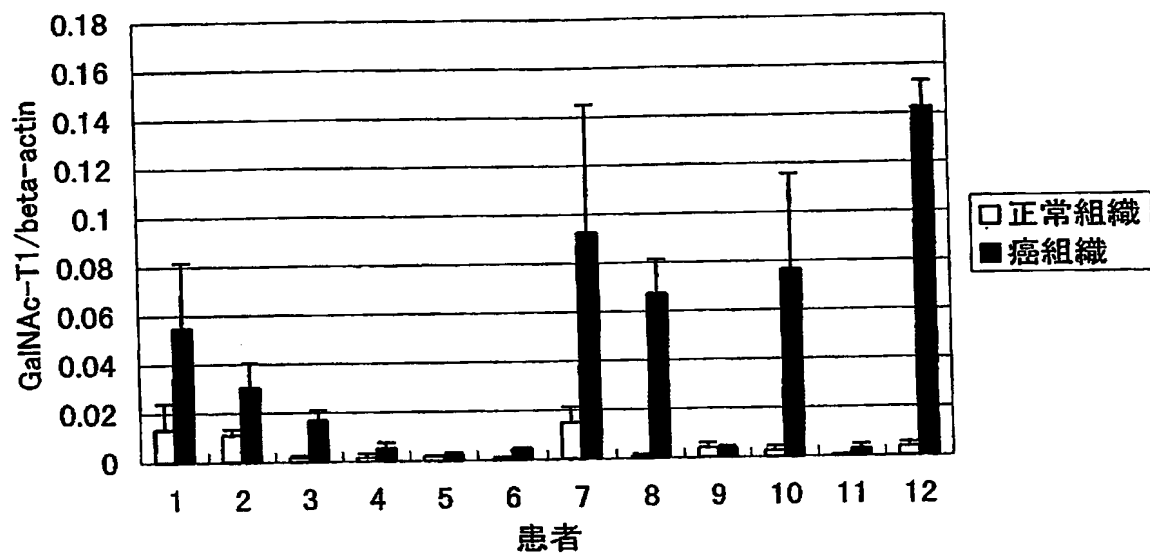
【図 1】



【図 2】

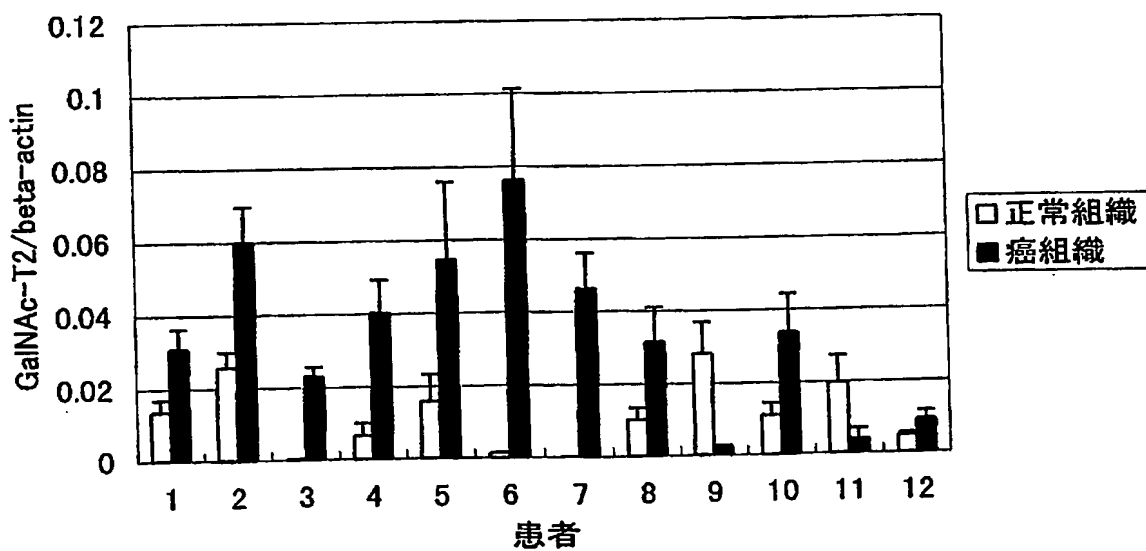
A

GalNAc-T1



B

GalNAc-T2



【書類名】 要約書

【要約】

【課題】 N-アセチルグルコサミンにN-アセチルガラクトサミンを転移する酵素が単離され、その遺伝子の構造が明らかにされた。したがって、該酵素等の遺伝子工学的な生産や該酵素による糖鎖の生産、該遺伝子らに基づく疾患の診断を可能にすること。

【解決手段】 配列表の配列番号1若しくは2若しくは3若しくは4に示されるアミノ酸配列または該アミノ配列において1若しくは複数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ配列に1若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有し、基質となるN-アセチルグルコサミンにN-アセチルガラクトサミン (GalNAc) を転移する活性を有するタンパク質および該タンパク質をコードする核酸を用いる。

【選択図】 なし

特願 2 0 0 2 - 2 3 6 2 9 2

出 願 人 履 歴 情 報

識別番号

[3 0 1 0 2 1 5 3 3]

1. 変更年月日
[変更理由]

住 所
氏 名

2 0 0 1 年 4 月 2 日

新規登録

東京都千代田区霞が関 1 - 3 - 1
独立行政法人産業技術総合研究所

特願 2002-236292

出願人履歴情報

識別番号

[597145779]

1. 変更年月日

2002年 1月25日

[変更理由]

名称変更

住 所

東京都新宿区百人町3丁目25番1号 サンケンビルディング

氏 名

アマシャム バイオサイエンス株式会社